VISIT…

LANZAROTE
Caliente.com

Г. Хакен

Информация и самоорганизация

Макроскопический подход к сложным системам



Издательство «Мир»

Информация и самоорганизация

Hermann Haken

Information and Self-Organization
A Macroscopic Approach to Complex Systems

Springer-Verlag Berlin Heidelberg
New York London Paris Tokyo

Г. Хакен

Информация и самоорганизация

Макроскопический подход к сложным системам

Перевод с английского
Ю. А. Данилова



Москва «Мир» 1991

ББК 22.317
X16
УДК 536

Хакен Г.

X16 Информация и самоорганизация: Макроскопический подход к сложным системам: Пер. с англ.–М.: Мир, 1991.–240 с., ил.

ISBN 5-03-001913-8

В книге автора из ФРГ развит оригинальный подход к описанию сложных макроскопических систем. Основное внимание уделено двум проблемам – использованию принципа максимума энтропии для неравновесных систем и распознаванию образов. Качественно новый уровень рассмотрения приближает реализацию физических систем нового типа (синергетический компьютер). Книга изобилует примерами из физики, химии, биологии, экономики, психологии.

Для специалистов – физиков, химиков, математиков, а также для аспирантов и студентов.

X $\frac{1604030000-062}{041(01)-91}$ 30–91 **ББК 22.317**

Редакция литературы по физике и астрономии

Научное издание

Герман Хакен

ИНФОРМАЦИЯ И САМООРГАНИЗАЦИЯ

Заведующий редакцией проф. А. Н. Матвеев. Зам. зав. редакцией С. М. Жебровский. Научный редактор Л. И. Третьякова. Мл. научный редактор В. И. Аксенова. Художник С. М. Болоболов. Художественный редактор К. В. Радченко. Технические редакторы О. Г. Лапко, Т. А. Мирошина. Корректор Т. М. Подгорная

ИБ № 7503

Сдано в набор 03.07.90. Подписано к печати 07.03.91. Формат 60 × 90 1/16. Бумага офсетная № 1. Печать офсетная. Гарнитура таймс. Объем 7,50 бум. л. Усл. печ. л. 15. Усл. кр.-отт. 30,42. Уч.-изд. л. 13,52. Изд. № 2/7659. Тираж 5000 экз. Зак. 890. Цена 3 р. 60 к.

Издательство «Мир» В/О «Совэкспорткнига» Государственного комитета СССР по печати. 129820, ГСП, Москва, И-110, 1-й Рижский пер., 2.

Можайский полиграфкомбинат В/О «Совэкспорткнига» Государственного комитета СССР по печати. г. Можайск, ул. Мира, 93.

ISBN 5-03-001913-8 (русск.)
ISBN 0-387-18639-5 (англ.)

© Springer-Verlag Berlin Heidelberg 1988
All Rights Reserved. Authorized translation from English language edition published by Springer-Verlag Berlin Heidelberg New York Tokyo
© перевод на русский язык, Данилов Ю. А., 1991

Предисловие проф. Ю. Л. Климонтовича

Профессор Г. Хакен хорошо известен многим читателям по многочисленным оригинальным работам и по книгам, большая часть которых издана в переводе на русский язык. И все же стоит, имея в виду прежде всего молодых читателей, сказать о нем несколько слов.

Г. Хакен начал свою научную карьеру в период бурного развития физики твердого тела, в период создания первой микроскопической теории сверхпроводимости и бурного развития теории полупроводников. Он является одним из пионеров создания квантовополевой теории экситонов и поляронов в твердых телах. Работы тех лет подытожены в его прекрасной книге «Квантовополевая теория твердого тела» (М.: Наука, 1980).

Одним из первых Г. Хакен включился и в создание теории лазеров. Он предложил, пожалуй, простейшую модель для описания когерентного лазерного излучения. Ему принадлежат работы по теории флуктуаций лазерного излучения. Так, например, хорошо известна формула Хакена–Лэкса для ширины линии лазерного излучения, в которой учтены тепловые флуктуации в резонаторе и спонтанное излучение атомов рабочей среды лазера.

Все также одним из первых Г. Хакен обратил внимание на глубокую аналогию процесса возникновения когерентного лазерного излучения с процессами, происходящими при фазовых переходах второго рода, на аналогию уравнений теории лазеров с уравнением Гинзбурга–Ландау в теории сверхпроводимости. В результате возник ставший теперь общепринятым термин «неравновесные фазовые переходы».

Именно в этот период деятельности проф. Г. Хакена возникла и штутгартская школа молодых физиков-теоретиков, которые в настоящее время занимают многие ведущие позиции в университетах ФРГ и определяют развитие ряда фундаментальных направлений современной статистической физики.

Исследования по теории лазеров и неравновесным фазовым переходам послужили началом нового объединяющего междисциплинарного направления, которое Г. Хакен обозначил весьма емким термином «синергетика». Этот термин подчеркивает принципиальную роль коллективных, кооперативных взаимодействий в возникновении и поддержании процессов самоорганизации в различных открытых системах.

Возникновение теории самоорганизации–синергетики–было подготовлено трудами многих выдающихся ученых. Это в первую очередь Ч. Дарвин–создатель теории биологической эволюции, Л. Больцман и А. Пуанкаре–основоположники статистического и

динамического описания сложных движений, а также А. Н. Колмогоров, Л. И. Мандельштам, А. А. Андронов, Н. С. Крылов, Н. М. Крылов и Н. Н. Боголюбов, А. А. Власов, Л. Д. Ландау, Я. Б. Зельдович и многие, многие другие.

Существенную роль в становлении теории самоорганизации сыграли работы Владимира Ивановича Вернадского о ноосфере. Созданию теории самоорганизации в современном ее понимании мы во многом обязаны И. Пригожину и Г. Хакену.

Напомним читателям, что Г. Хакен является инициатором выпуска издательством «Шпрингер» специальной серии книг по синергетике. Ее с полным правом можно назвать интернациональной. Две книги этой серии (Синергетика, М.: Мир, 1980 и 1985) принадлежат Г. Хакену. Первая из них открывает серию и может служить введением в синергетику. Вторая – имеет подзаголовок «Иерархия неустойчивостей в самоорганизующихся системах и устройствах» и, таким образом, посвящена более конкретным аспектам синергетики.

Предлагаемая читателям книга Г. Хакена «Информация и самоорганизация» – уже 40-й том шпрингеровской серии книг по синергетике. Для облегчения понимания роли этой книги на фоне современного развития теории самоорганизации отметим следующее.

Название книги сопровождается подзаголовком «Макроскопический подход к сложным системам». Автор подчеркивает тем самым, что речь идет о проблемах, которые решаются на уровне макроскопического описания. Это одновременно и недостаток, поскольку в принципе более глубокой является микроскопическая теория, и преимущество, поскольку для решения практических задач макроскопический (феноменологический) подход во многих случаях оказывается более эффективным,– но при обязательном условии достаточно развитой физической интуиции. Это отмечается и автором выбором названия первой главы: «Сложные системы – вызов искусству исследователя». Трудности начинаются уже с попытки дать исчерпывающее определение фундаментальных понятий: «сложность», «самоорганизация». Такие определения, однако, едва ли необходимы, поскольку основную ценность представляют не абсолютные определения, а сопоставление значений относительной степени сложности и упорядоченности различных состояний открытых систем. Это необходимо, в частности, и для констатации того, является ли рассматриваемый процесс процессом самоорганизации или, напротив, деградации.

Мы еще вернемся к этим вопросам. Здесь же лишь отметим, что вводная глава написана с большим мастерством и служит ключом к решению комплекса проблем, который автор закодировал названием «информация и самоорганизация».

Глава 2 «От микроскопического мира к макроскопическому» служит первой ступенькой на пути к решению поставленных во введении задач. Автор выделяет три уровня описания: микроскопи-

ческий, мезоскопический и макроскопический. На примере теории газов этому отвечают соответственно описание на основе обратимых уравнений движения атомов – уравнений Гамильтона, на основе кинетического уравнения Больцмана и на основе уравнений газовой динамики. Второй и третий уровни описания опираются на представления о газе как о сплошной среде и не дают полного описания системы (полного в микроскопическом смысле). При этом (чаще неявно) проводится осреднение микроскопических характеристик по бесконечно малому (в физическом смысле) объему или временному интервалу, что с неизбежностью приводит к необратимым уравнениям.

В соответствии с задачами книги в этой главе речь идет о соотношении мезоскопического (кинетического в теории газов) и макроскопического описаний. В кинетической теории газов мезоскопическими являются координаты и скорости малых, но макроскопических элементов среды, функция распределения которых находится путем решения уравнения Больцмана. В табл. 2.1 приведены примеры мезоскопических переменных, рассматриваемых в книге. Соответствующие функции распределения удовлетворяют уравнениям Фоккера – Планка, которые могут быть заменены эквивалентными уравнениями Ланжевена – стохастическими уравнениями. Здесь необходимо одно замечание, связанное с рассматриваемой кратко и в этой книге «дилеммой Ито – Стратоновича».

Вопрос связан с неоднозначностью (при заданных нелинейных динамических уравнениях рассматриваемой системы) формулировки уравнений Фоккера – Планка и соответствующих уравнений Ланжевена. Оказывается, что одному и тому же нелинейному динамическому уравнению отвечают два различных в общем случае уравнения Фоккера – Планка и уравнения Ланжевена. Возникает естественно вопрос о том, какому из этих двух уравнений отдать предпочтение. Существенно, что есть и еще одна выделенная форма уравнения Фоккера – Планка для нелинейных систем, которую можно назвать «кинетическая форма» (см. [2, 4]). В последнем случае формула Эйнштейна, связывающая коэффициент диффузии (интенсивность источника Ланжевена) и коэффициент трения динамической системы, остается справедливой и для нелинейных систем. В [2, 5] приведена аргументация в пользу «кинетической формы» записи уравнений Фоккера – Планка. Это надо иметь в виду, поскольку различие решений указанных трех типов уравнений Фоккера – Планка может оказаться существенным.

Глава 3 подводит читателя вплотную к центральной теме книги. Здесь сформулирован принцип максимума энтропии. В наиболее общем виде он был сформулирован, по-видимому, в работах Джейнса. При этом в качестве дополнительных условий задавались значения сохраняющихся величин.

Возможна, однако, и значительно более общая трактовка этого

принципа, позволяющая использовать его для широкого круга неравновесных процессов. В связи с этим надо отметить ставшие уже классическими работы Д. Н. Зубарева [6], в которых развит метод неравновесного статистического оператора – обобщенного на неравновесные процессы канонического распределения Гиббса. При этом в качестве дополнительных условий используются не постоянные величины, а функции, удовлетворяющие уравнениям, выражающим законы сохранения. Примером могут служить гидродинамические функции.

В данной книге Г. Хакен сделал существенный новый шаг в применении принципа максимума энтропии. Именно, в число дополнительных условий включены и функции, играющие роль параметров порядка при неравновесных фазовых переходах. Тем самым принцип максимума энтропии становится рабочим инструментом в теории самоорганизации – синергетике.

Макроскопический принцип максимума энтропии используется в гл. 5 как пробный камень для нахождения стационарных распределений переменных электромагнитного поля в одномодовом и многомодовом лазерах. В результате получаются известные выражения для стационарных распределений. Рассмотрен также и менее традиционный пример.

В гл. 6 исследуется вопрос о выделении распределений параметров порядка, когда в качестве дополнительных условий задаются первые четыре момента. Обсуждаются также возможности использования принципа максимума энтропии в теории неравновесных фазовых переходов.

Глава 7 – центральная как по положению в книге, так и по существу вопроса. Действительно, именно здесь сталкиваются два основных понятия: информация и самоорганизация, которые вынесены в основной заголовок книги. В этой главе, в частности, проведен подробный анализ изменения информации (энтропии Шеннона) в самоорганизующейся системе при изменении управляющего (контрольного) параметра. Иными словами, проведен анализ изменения информации Шеннона при неравновесном фазовом переходе в состояние с меньшей симметрией, т. е. в состояние с большей степенью упорядоченности. Показано, что при таком неравновесном фазовом переходе система приобретает способность хранить информацию. Это дает автору основание трактовать энтропию Шеннона как информацию, т. е. предпочесть термину «энтропия» термин «информация».

Однако, и автор сразу же это отмечает, в области точки перехода (вблизи порога) из-за критических флуктуаций информация сильно возрастает, поэтому в данной области термин «энтропия» оказывается предпочтительней.

С другой стороны, при использовании энтропии Шеннона в качестве меры упорядоченности возникает еще одна «трудность».

Именно, при переходе, например, в лазере через порог генерации лазерного излучения энтропия Шеннона оказывается больше, чем в исходном «равновесном» состоянии. В такой ситуации трактовка энтропии Шеннона как информации снова оказывается предпочтительней, поскольку представление состояния развитой генерации как более хаотического противоречит здравому смыслу – противоречит физической интуиции.

Выход из этого «трудного» положения оказывается возможным, если расчет энтропии производить по рецепту S-теоремы (см. разд. 7.10 книги), т. е. относительную степень упорядоченности оценивать по значениям энтропии при заданном значении средней эффективной энергии состояний рассматриваемой открытой системы. Чтобы изложенный в разд. 7.10 материал был более понятен, необходимо кратко сформулировать сущность утверждения, названного S-теоремой. Подробное изложение этого вопроса можно найти в работах [2, 7].

Следующие три главы развивают в разных аспектах результаты, отмеченные выше. Особого внимания заслуживает гл. 9, в которой принцип максимума энтропии используется для конструирования зависящих от времени распределений. Исходным пунктом служит так называемый принцип максимального калибра, сформулированный Джейнсом. Имеется, несомненно, определенная связь этого подхода с упомянутым выше методом неравновесного статистического оператора. Сравнительный анализ этих двух подходов до настоящего времени, по-видимому, не проведен, поэтому отметим лишь эффективность реализованного в книге Г. Хакена подхода для конструирования уравнений, на основе которых можно проследить за временным ходом процессов самоорганизации.

Очень изящна и увлекательна гл. 11, в которой изложенные в предыдущих главах методы применяются к биологическим системам. Этим демонстрируется эффективность построения теории неравновесных фазовых переходов на основе принципа максимума энтропии. Здесь речь идет о смене типа движения у человека и животных при увеличении частоты двигательной активности; точке неравновесного фазового перехода отвечает критическое значение частоты.

Поражает насыщенностью гл. 12. В ней автор излагает основы предложенного им синергетического подхода к проблеме распознавания образов.

Предложенный Г. Хакеном подход к распознаванию образов был проверен с помощью численного эксперимента. Для этого были подготовлены фотографии нескольких человек. Их фамилии были закодированы в правом верхнем углу фотографии. Затем компьютеру предъявлялись портреты или их части. При этом компьютер восстанавливал полностью лицо и фамилию и тем самым распознавал образ.

В этой же главе, все так же основываясь на принципе максимума

энтропии, Г. Хакен формулирует основные принципы предлагаемого им синергетического компьютера.

В заключительной части главы показано, каким образом может быть восстановлена физическая система, воспроизводящая марковский процесс. Существенно, что рассматриваемая система, по-видимому, способна не только воспроизводить марковский процесс по начальным данным, – она может и обладать ассоциативной памятью, т. е. восстанавливать процесс по неполным данным.

Следующая глава посвящена проблемам информации и самоорганизации в квантовой теории. Здесь снова выступает принцип максимума энтропии, на основе которого восстанавливается вид уравнения для матрицы плотности. Автор, однако, сосредоточивает свое внимание на моментах распределений. Он показывает, что многие результаты, полученные в рамках классической теории, допускают квантовое обобщение.

В краткой заключительной главе автор обращает внимание читателя на те области, в которых остаются еще нерешенные вопросы.

Итак, мы завершаем путешествие по страницам этой небольшой, но удивительно содержательной книги. Нет оснований утверждать, что все изложенное представляет собой завершенную картину области знания, которую автор обозначил «информация и самоорганизация». Это все же скорее первые наброски, но они в большой степени дают представление об облике будущего здания. Существенно, что ценность книги состоит не только в изложении конкретных результатов, но и в значительной степени в том, что она стимулирует дальнейшие исследования в этой увлекательной и необычайно важной в практическом отношении области знания.

Ю. Л. Климонтович

Литература

1. *Анищенко В. С.* Сложные колебания в простых системах: Механизмы возникновения, структуры и свойства динамического хаоса в радиофизических системах. – М.: Наука, 1990.
2. *Климонтович Ю. Л.* Турбулентное движение и структура хаоса: Новый подход к статистической теории открытых систем. – М.: Наука, 1990.
3. *Лоскутов А. Ю., Михайлов А. С.* Введение в синергетику. – М.: Наука, 1990.
4. *Лифшиц Е. М., Питаевский Л. П.* Физическая кинетика. – М.: Наука, 1979.
5. *Klimontovich Yu. L.* Ito, Stratonovich Kinetic Forms of the Stochastic Equations, Physica A, 1990.
6. *Зубарев Д. Н.* Неравновесная статистическая термодинамика. – М.: Наука, 1971.
7. *Климонтович Ю. Л.* Проблемы статистической теории открытых систем: критерии относительной степени упорядоченности состояний в процессах самоорганизации. – УФН, 1989, т. 158, с. 59.
8. *Стратонович Р. Л.* Теория информации. – М.: Советское радио, 1985.
9. *Стратонович Р. Л.* Нелинейная неравновесная термодинамика. – М.: Наука, 1985.

Предисловие к русскому изданию

Согласно современным научным представлениям, все живые существа обретают форму и функции с помощью самоорганизации. За последние двадцать лет или около того стало ясно, что самоорганизация происходит и в неорганическом мире, изучением которого занимается физика и химия. Широко известным тому примером может служить реакция Белоусова–Жаботинского, открытая в Советском Союзе. В двух моих предыдущих книгах, а именно «Синергетика» и «Синергетика. Иерархии неустойчивостей», были изложены принципы, позволяющие в рамках единого подхода рассматривать широкий класс явлений самоорганизации, происходящих как в мире живого, так и в неорганическом мире. Эти принципы опирались на микроскопическую или мезоскопическую теорию.

В книге, предлагаемой вниманию читателя, я показываю, что возможен и макроскопический подход к теории процессов самоорганизации, и в качестве исходного пункта я принимаю принцип максимума информационной энтропии. Это приводит нас к центральному вопросу о той роли, которую энтропия играет в сильно неравновесных процессах, – вопросу, в решение которого фундаментальный вклад внесли работы проф. Ю. Л. Климонтовича.

Выход этой книги в русском переводе для меня высокая честь, и я хочу поблагодарить проф. Климонтовича за все те усилия, которые он приложил к изданию этой книги. Я надеюсь, что профессора, студенты и исследователи, работающие в различных областях науки, найдут в этой книге немало полезного для своих собственных исследований.

Штутгарт, июль 1989 *Г. Хакен*

Памяти Эдит

Предисловие

Сложные системы встречаются повсюду и служат предметом изучения практически всех областей науки от физики через химию и биологию до экономики и социологии. В этой книге мы намереваемся изложить понятия и методы, позволяющие рассматривать сложные системы с единой точки зрения. Именно поэтому наша книга может представить интерес для аспирантов, профессоров и научных работников, ведущих теоретические исследования в указанных выше областях науки. Основная идея нашего единого подхода восходит к основной идее синергетики. Чтобы найти единые принципы, мы сосредоточим внимание на таких ситуациях, когда сложная система качественно изменяет свое макроскопическое поведение, или, иначе говоря, когда она изменяет свою макроскопическую пространственную, временную или функциональную структуру.

До сих пор синергетика как теория обычно начиналась с микроскопического или мезоскопического описания сложной системы. В этой книге мы излагаем подход, который начинается с макроскопических данных. В частности, мы рассмотрим системы, обретающие свою новую структуру без особого воздействия извне, т.е. самоорганизующиеся системы. Инструментом нашего анализа будет информация. Так как это слово имеет несколько совершенно различных значений и все эти значения важны для наших целей, мы обсудим различные аспекты информации. Диапазон их необычайно широк – от шенноновской информации, напрочь лишенной какой бы то ни было семантики, до воздействия информации на тех, кто ее принимает, и самозарождения значения.

Шенноновская информация тесно связана с введенной Больцманом статистической энтропией. Совершенно общая формулировка была дана Джейнсом в форме принципа максимума энтропии, который по причинам, излагаемым в нашей книге, мы будем называть *принципом максимума информационной энтропии*. Как было показано Джейнсом, этот принцип позволяет весьма изящно выводить основные соотношения термодинамики и может быть положен в основу подхода к неравновесной термодинамике. Ингарден сформулировал то, что он назвал *информационной термодинамикой*, вводя температуры более высоких порядков. Несмотря на успех, принцип максимума информационной энтропии был подвергнут критике за субъективность, поскольку выбор ограничений, при которых находится максимум энтропии, представляется произвольным. Но с помощью результатов синергетики эту проблему удается решить для широкого класса явлений, а именно для самоорганизующихся систем, в которых новая структура образуется посредством *неравновесного фазового перехода*.

Таким образом, наш подход применим к многим наиболее интересным ситуациям. Мы будем иллюстрировать наш общий подход примерами из физики (лазеры, гидродинамика), компьютерных наук (распознавание образов с помощью машин) и биологии (морфогенез поведения). Последний пример особенно наглядно и убедительно демонстрирует возможность применения нашего подхода к действительно сложным системам и показывает, что поведение таких систем может быть смоделировано с помощью четко определенной процедуры.

Тем самым наша книга знакомит читателя и с общей теорией, и с практическими приложениями. Я надеюсь, что она окажется полезной и для учебных целей, и в исследованиях сложных систем. Выражаю свою признательность проф. Х. Шимидзу за полезные обсуждения разд. 1.6 и моим сотрудникам д-ру В. Банцхофу, М. Бестехорну, В. Лоренцу, М. Шинделю и Ф. Веберуссу за помощь в чтении корректур. Я хотел бы поблагодарить г-жу А. Конц и г-жу И. Мёллер, безукоризненно перепечатавших несколько вариантов рукописи, и г-на А. Фукса и г-на В. Лоренца за подготовку иллюстраций.

Считаю своим приятным долгом поблагодарить Ганноверский фонд фирмы «Фольксваген» за финансовую поддержку проекта «Синергетика». Наконец, хотя и далеко не в последнюю очередь, я хотел бы выразить свою признательность издательству «Шпрингер», в особенности д-ру Анжеле Лехи и д-ру Хельмуту Лотшу, за готовность к сотрудничеству.

Штутгарт, январь 1988 *Г. Хакен*

Глава 1
Сложные системы – вызов искусству исследователя

Цель этой книги состоит в том, чтобы изложить понятия и методы, которые позволяют нам подходить к рассмотрению сложных систем с единой точки зрения. Наша книга состоит из двух частей: в вводной главе теория сложных систем строится на качественном уровне, в то время как остальная часть книги посвящена количественным методам. В гл. 1 мы познакомимся с примерами сложных систем и некоторыми типичными подходами к их рассмотрению, в частности с термодинамикой и синергетикой. Мы обсудим понятие самоорганизации и, в частности, различные аспекты информации. В заключительном разделе этой главы дается сжатый обзор нашей новой теории, которую можно рассматривать как *макроскопический* подход к синергетике. *Микроскопический* подход к синергетике кратко излагается в гл. 2, а гл. 3 знакомит читателя с введением в принцип максимума информационной энтропии, играющей важную роль в нашей книге. Гл. 4 иллюстрирует этот принцип на примере его приложения к термодинамике.

Остальная часть этой книги посвящена нашему количественному методу и его приложениям; приводятся подробные примеры из физики и биологии. Наконец, будет показано, что один важный подход в области распознавания образов содержится в качестве частного случая в нашей общей теории. Тем самым наш подход позволяет достичь существенной унификации в науке. Читатели, для которых качественное рассмотрение сложных систем не представляет особого интереса, могут опустить эту вводную главу и перейти в зависимости от уровня своих знаний прямо к гл. 2 или 3 и 4, или 5.

Мы же начнем с некоторых основных понятий.

1.1. Что такое сложные системы?

Прежде всего нам необходимо обсудить, чтó мы понимаем под сложными системами. При наивном подходе их можно описать как системы, состоящие из большого числа частей, элементов или компонентов, которые могут быть как одного, так и различного рода. Компоненты, или части, могут соединяться между собой более или менее сложным образом. Различные области науки дают нам многочисленные примеры систем, называемых сложными. Одни из них оказываются весьма простыми, в то время как другие оказываются действительно сложными.

Начнем с нескольких примеров из физики. Газ состоит из очень многих молекул, число их в кубическом сантиметре достигает 10^{22}. Молекулы газа летят совершенно беспорядочно и поэтому претерпевают друг с другом многочисленные столкновения (рис. 1.1). В отличие от молекул газа в кристалле атомы или молекулы расположены в высшей степени упорядоченно и совершают лишь слабые колебания (рис. 1.2). Нас могут интересовать какие-нибудь конкретные свойства, например давление или температура газа или сжимаемость кристалла. Мы можем рассматривать сложные системы и с точки зрения их практической пользы; например, газ и водяной пар могут использоваться как рабочее тело в паровом двигателе, кристалл может найти применение как проводник электричества и т. д. Некоторые физические системы были созданы в основном для достижения вполне определенной цели. К числу таких физических систем относится, например, лазер (рис. 1.3). Этот новый источник света был создан для получения излучения особого типа.

В химии мы также встречаемся со сложными системами. В химических реакциях участвует очень много молекул и образуются новые молекулы. Изобилует сложными системами биология. Клетка со-

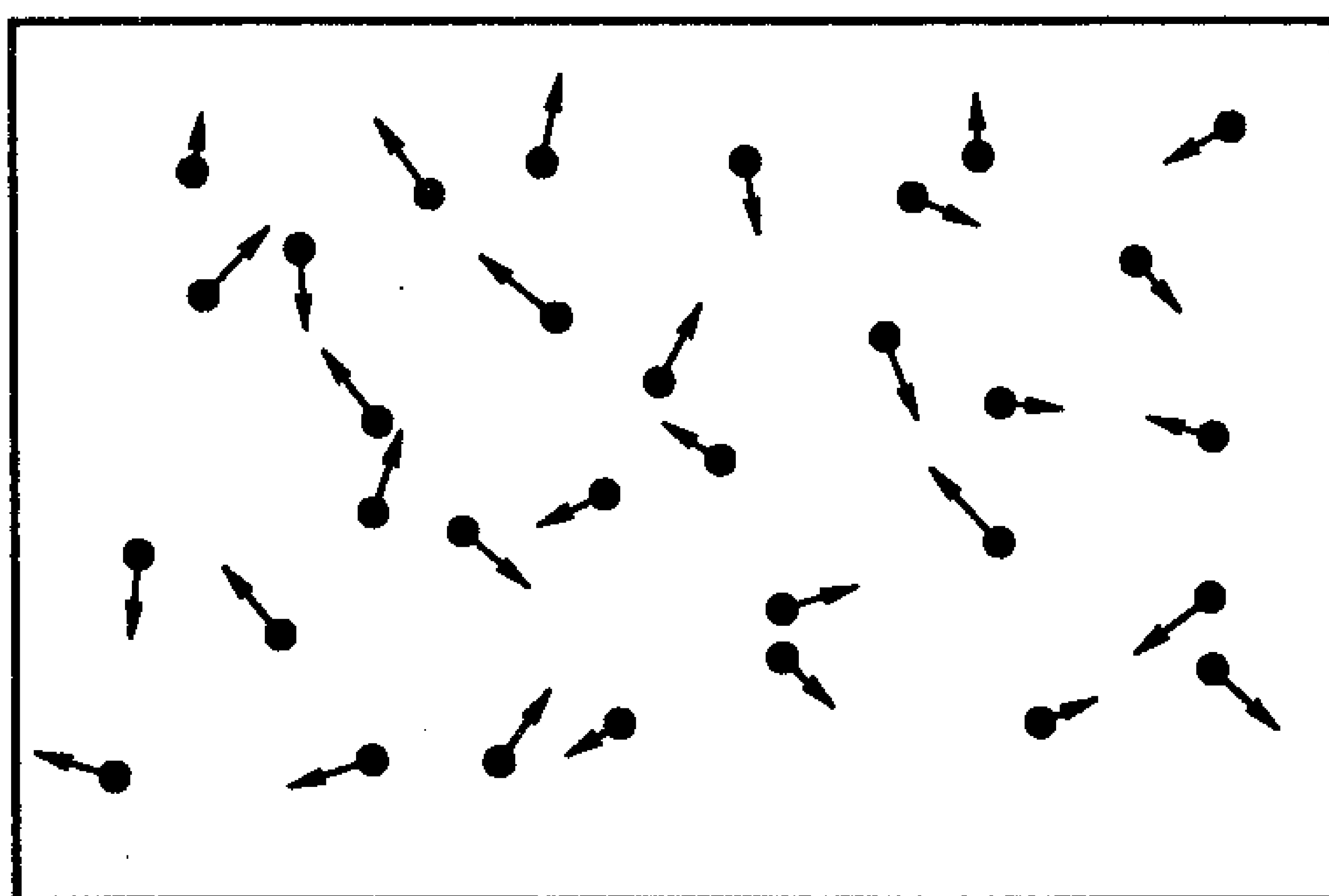


Рис. 1.1. Атомы газа, движущиеся в ящике.

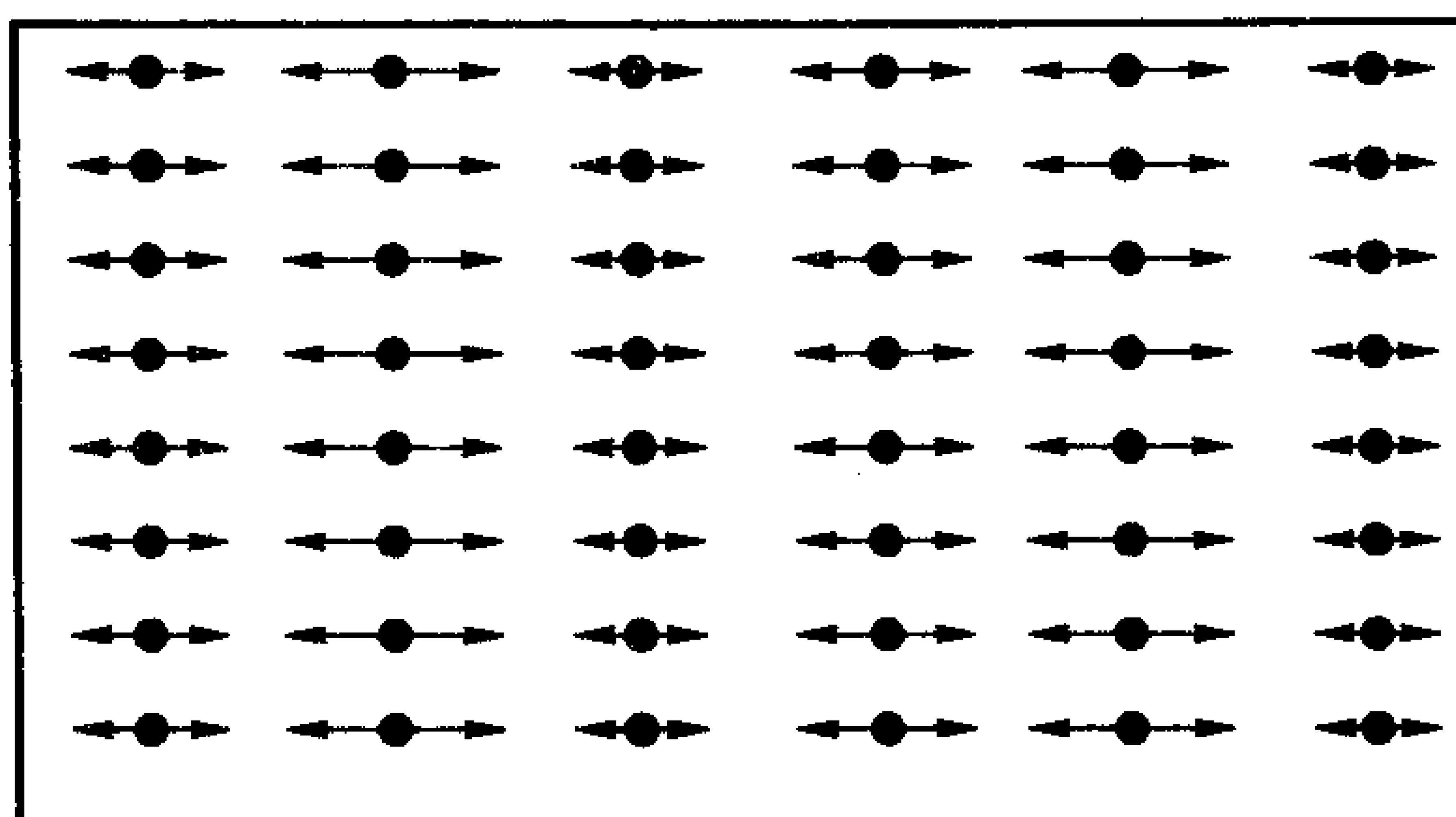


Рис. 1.2. Атомы в кристалле.

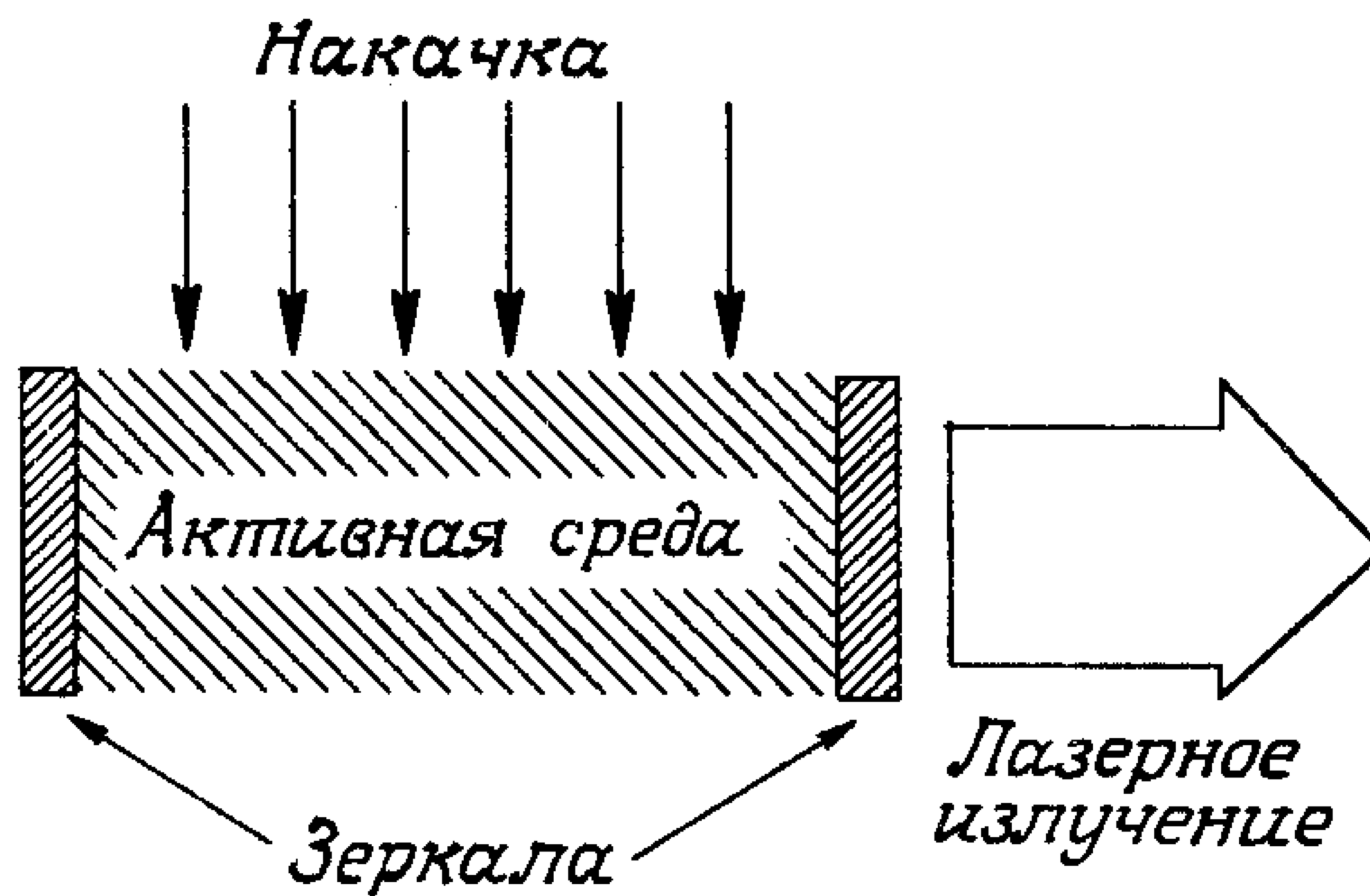


Рис. 1.3. Схема действия лазера.

стоит из клеточной мембраны, имеющей весьма непростую структуру, ядра и цитоплазмы, которые в свою очередь состоят из многих и многих компонентов (рис. 1.4). В клетке многие тысячи метаболических процессов могут протекать одновременно и согласованно. Органы тела человека и животных состоят из огромного числа клеток, которые взаимодействуют между собой вполне согласованно. В свою очередь органы выполняют специфические функции и взаимодействуют между собой в организме животного. Животные образуют сообщества (рис. 1.5). По-видимому, наиболее сложной системой в мире является головной мозг человека, состоящий из 10^{10} или более нервных клеток (рис. 1.6). Согласованное действие этих клеток позволяет нам распознавать образы, говорить и выполнять другие интеллектуальные функции.

В технике мы также встречаемся с сложными системами. Такие системы могут быть машинами, например двигателем автомобиля, целыми заводами или электростанциями, соединенными в энергети-

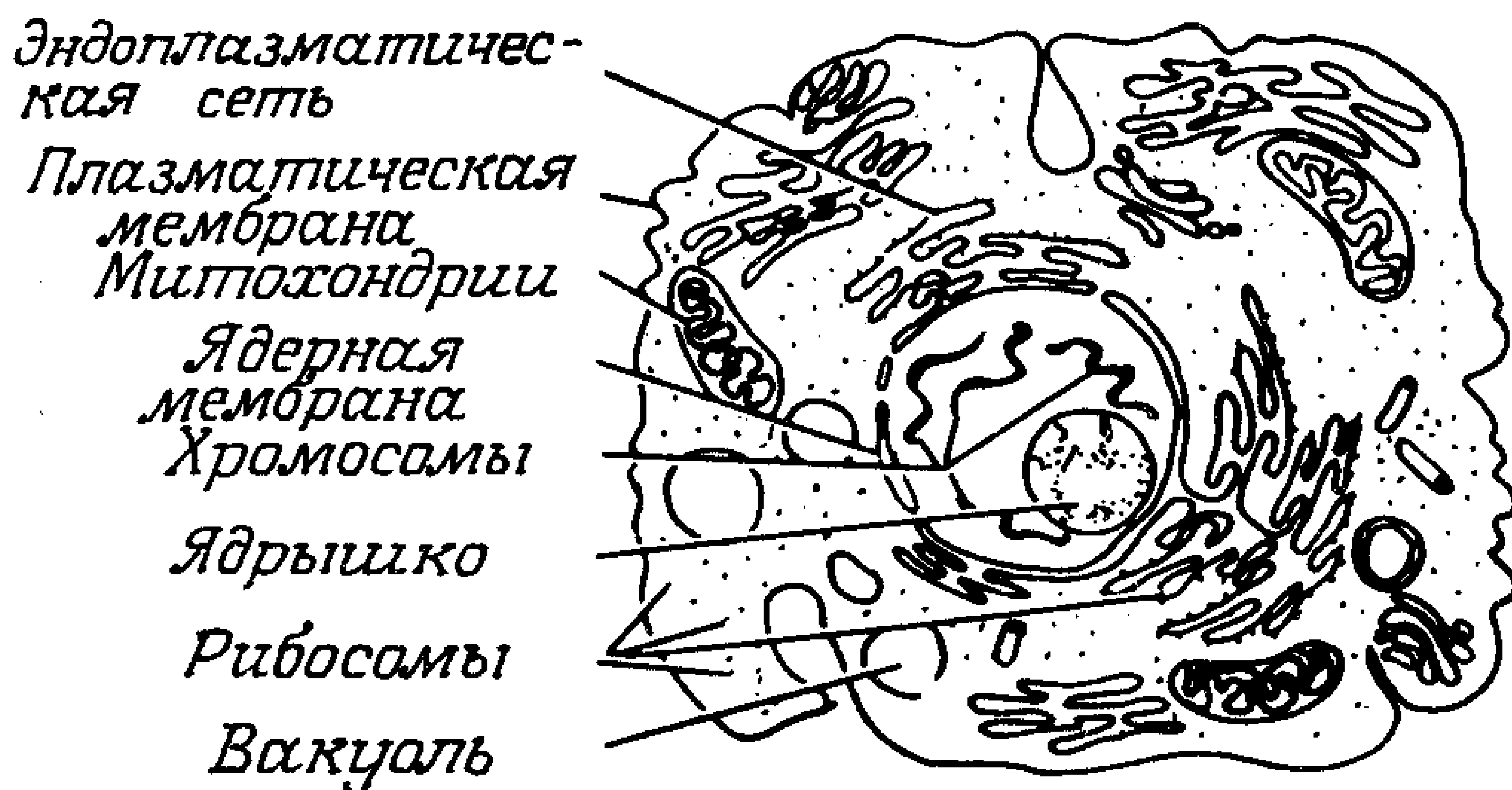


Рис. 1.4. Типичная биологическая клетка. [Из книги: *Anderson D. A., Sobieski R. J.*, Introduction to Microbiology, C. V. Mosby Company, 1980.]

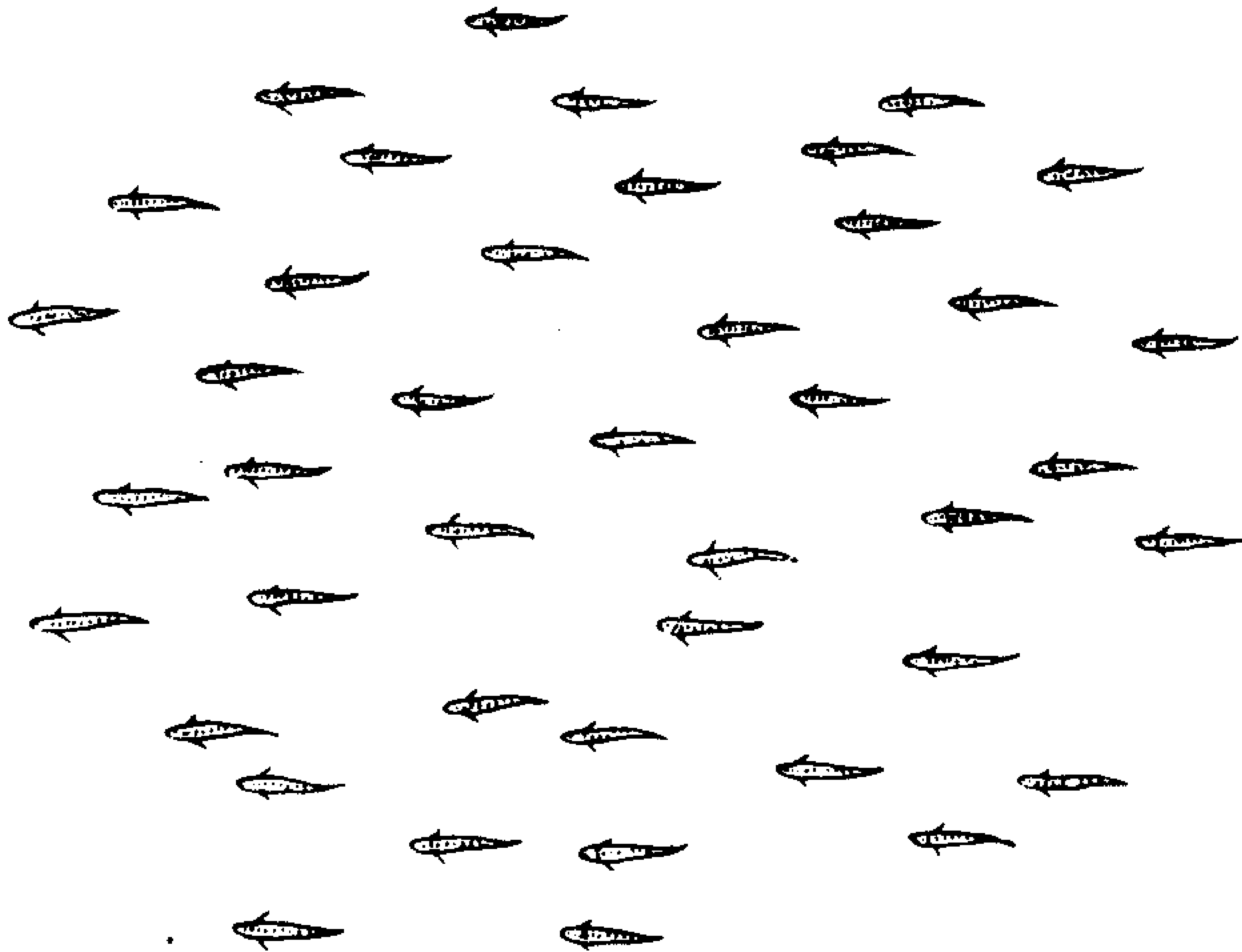


Рис. 1.5. Стая рыб. [Из работы: *Patridge B.L.*, Wie Fische Zusammenhalten. © Spectrum der Wissenschaft (Aug. 1982).]

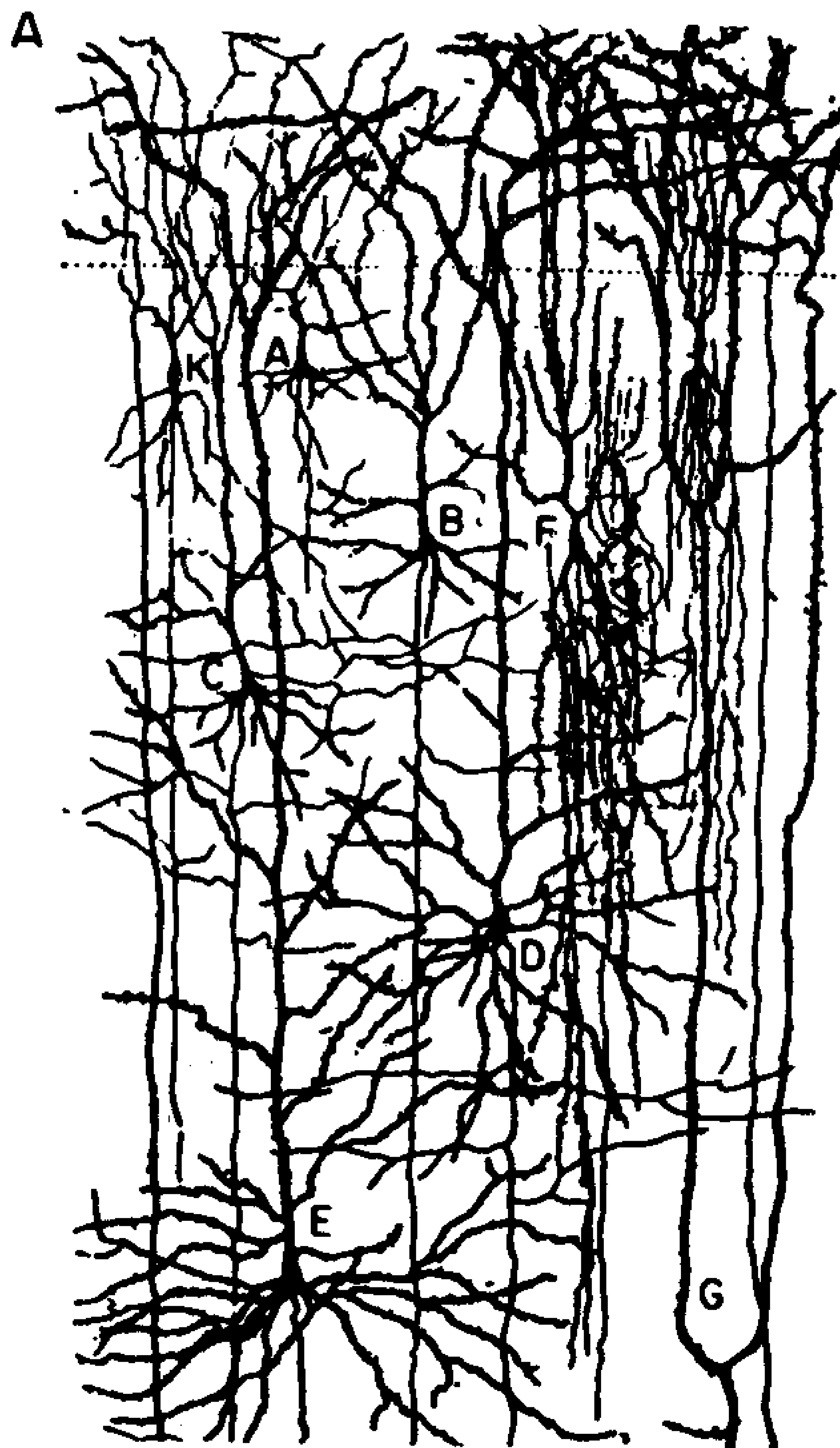


Рис. 1.6. Сеть нервных клеток. [Из книги: *Quarton G.C., Melnechuck T., Schmitt F.O.*, The Neuro-sciences, New York, The Rockefeller University Press, 1967.]

ческую сеть. Экономика с ее многочисленными участниками, товарооборотом и денежным обращением, транспортными потоками, производством, потреблением и хранением товаров может служить еще одним примером сложной системы. Аналогично и все общество в целом с многообразными видами человеческой деятельности, политическими, религиозными, профессиональными или культурными различиями людей является примером сложной системы. Компьютеры все более выходят на тот уровень, когда их можно будет считать сложными системами. В особенности это относится к компьютерам так называемого 5-го поколения, в которых обработка информации идет на смену перемалыванию чисел, столь характерному для современных компьютеров.

Системы могут быть сложными не только потому, что они состоят из большого числа частей: мы можем говорить и о сложном поведении. Так, очень сложными могут быть различные проявления человеческого поведения, например те, изучением которых занимается психология. С другой стороны, нас восхищает также высокая степень координации мышц при движении, дыхании и т. д. (рис. 1.7). Наконец, современная наука сама по себе является сложной системой, что становится совершенно ясным, если принять во внимание огромное число различных областей знания.

Уместно спросить, почему многие системы так сложны и как они возникли. Например, в технике и в биологии необходимость сложности вполне очевидна. Сложные системы предназначены для выполнения определенных функций, и, как показывает тщательный анализ, эти функции могут быть выполнены только сложной системой, состоящей из многих согласованно действующих частей. Когда же мы говорим о возникновении сложных систем, необходимо различать системы двух типов. С одной стороны, мы имеем системы, созданные человеком, спроектированные и построенные людьми. Это машины или конструкции, предназначенные для решения определенных задач. С другой стороны, существует очень много сложных систем в природе, созданных самой природой, или, иначе говоря, возникших в результате самоорганизации. Совершенно очевидно в этой связи, что в биологии важную роль играет эволюционный подход, или дарвинизм, представляющий собой попытку понять,

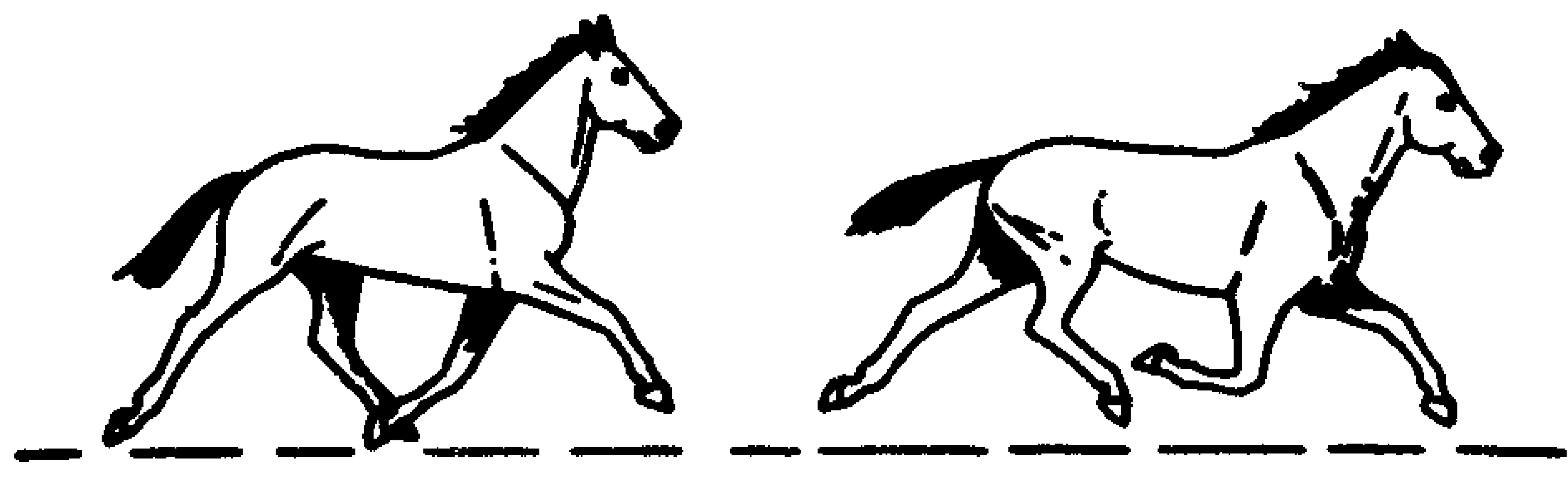


Рис. 1.7. Лошадь, бегущая рысью. [Из книги: *Kolb E.*, Lehrbuch der Physiologie der Haustiere, Stuttgart, Fischer-Verlag, 1967.]

почему и как в ходе эволюции возникают все более и более сложные системы.

После весьма поверхностного и беглого обзора сложных систем попытаемся теперь дать более строгое определение. Современное определение сложной системы опирается на понятие алгебраической сложности. По крайней мере в определенных пределах систему можно описать как строку, или последовательность, данных, например флуктуирующей интенсивности света, приходящего от звезд (рис. 1.8), или температурной кривой больного, где данные представлены числами. Итак, рассмотрим какую-нибудь последовательность чисел и попробуем определить сложность такой последовательности. Если иметь в виду конкретные примеры числовых последовательностей, например последовательность 1, 4, 9, 16, 25, 36, ..., то нетрудно понять, что такая последовательность может быть образована по простому закону, в нашем случае по закону n^2, где n – целое число. Следовательно, если нам задана строка данных, то позволительно спросить, существует ли компьютерная программа и множество начальных данных, по которым эта программа может вычислить всю строку данных. Разумеется, длина программы может варьироваться в зависимости от конструкции компьютера.

Следовательно, для того, чтобы мы могли сравнивать длину программ, нам необходим универсальный компьютер. Не вдаваясь в детали, можно утверждать, что такой универсальный компьютер может быть построен по крайней мере в мысленном эксперименте, как это было показано Тьюрингом (рис. 1.9). В честь чего мы будем называть такой универсальный компьютер машиной Тьюринга.

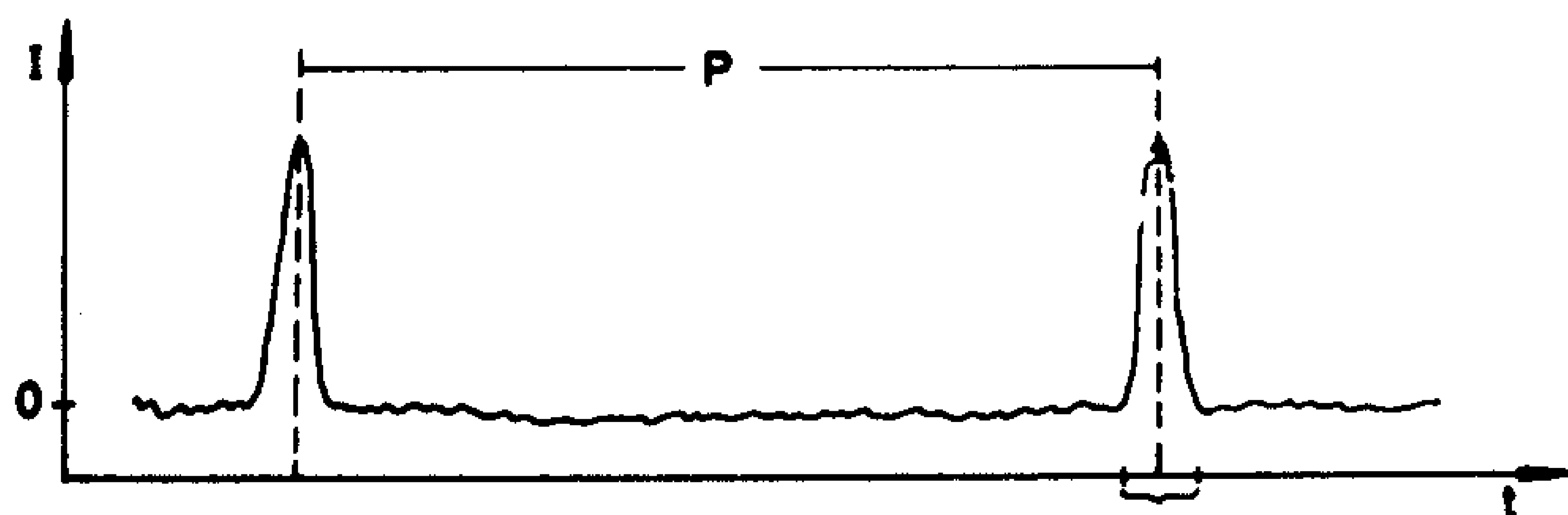


Рис. 1.8. Световой импульс от пульсара. [Из книги: *Weigert, Wendke,* Astronomie und Astrophysik, Weinheim, Physik-Verlag, 1982.]

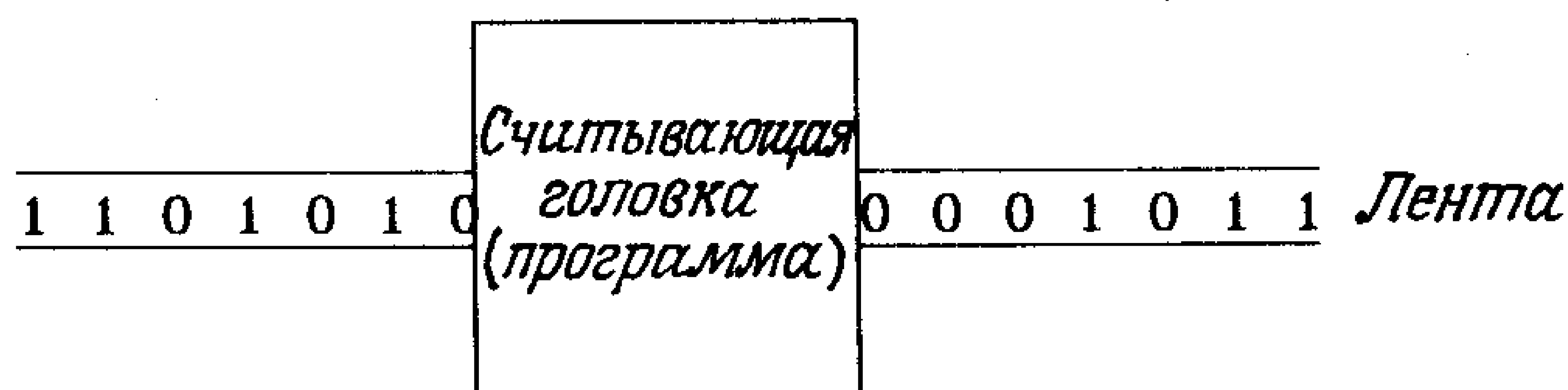


Рис. 1.9. Схема машины Тьюринга.

Далее основная идея состоит в том, чтобы сжать до минимума программу и начальное множество данных. Минимальная длина программы и множества начальных данных служит мерой алгебраической степени сложности (рис. 1.10). Однако у такого определения есть уязвимое место. Как показывает знаменитая теорема Гёделя, проблема нахождения минимальной программы и минимального объема начальных данных не имеет универсального решения. Иначе говоря, не существует общего алгоритма, который позволил бы решить эту проблему. Такого рода алгоритмы удается создавать только в отдельных частных случаях. Лишь изредка кому-нибудь приходит в голову остроумная идея, позволяющая неожиданно просто решить трудную задачу. Рассмотрим в качестве примера газ. Можно было попытаться проследить траектории отдельных частиц и их столкновения, а затем построить функцию распределения скоростей частиц. Если иметь в виду решение этой задачи в смысле макроскопического описания, то она так и не была решена. Но простой и изящный вывод функции распределения, известной под названием распределения Больцмана, удалось осуществить в рамках статистической механики, не прибегая к микроскопическому подходу, а используя понятие энтропии (см. ниже). Можно привести ряд аналогичных примеров, показывающих, что существуют оригинальные подходы, позволяющие найти неожиданно простое решение проблемы, первоначально казавшейся почти неприступной. Резюмируя, можно сказать, что мы отчетливо сознаем, сколь тонкое понятие сложность. Главная цель нашей книги как раз и состоит в том, чтобы предложить с единой точки зрения некоторые оригинальные подходы, позволяющие эффективно решать проблемы, связанные с сложными системами.

Сложные системы можно рассматривать с различных точек зрения. Например, биологическую систему можно рассматривать на макроскопическом уровне, исследуя ее поведение, или на промежуточном уровне, изучая функционирование ее органов, или, наконец, заняться исследованием химии ДНК. Объем данных, которые иногда бывает необходимо собирать тем, кто занимается изучением сложных систем, часто оказывается необъятно большим. Кроме того, выбрать

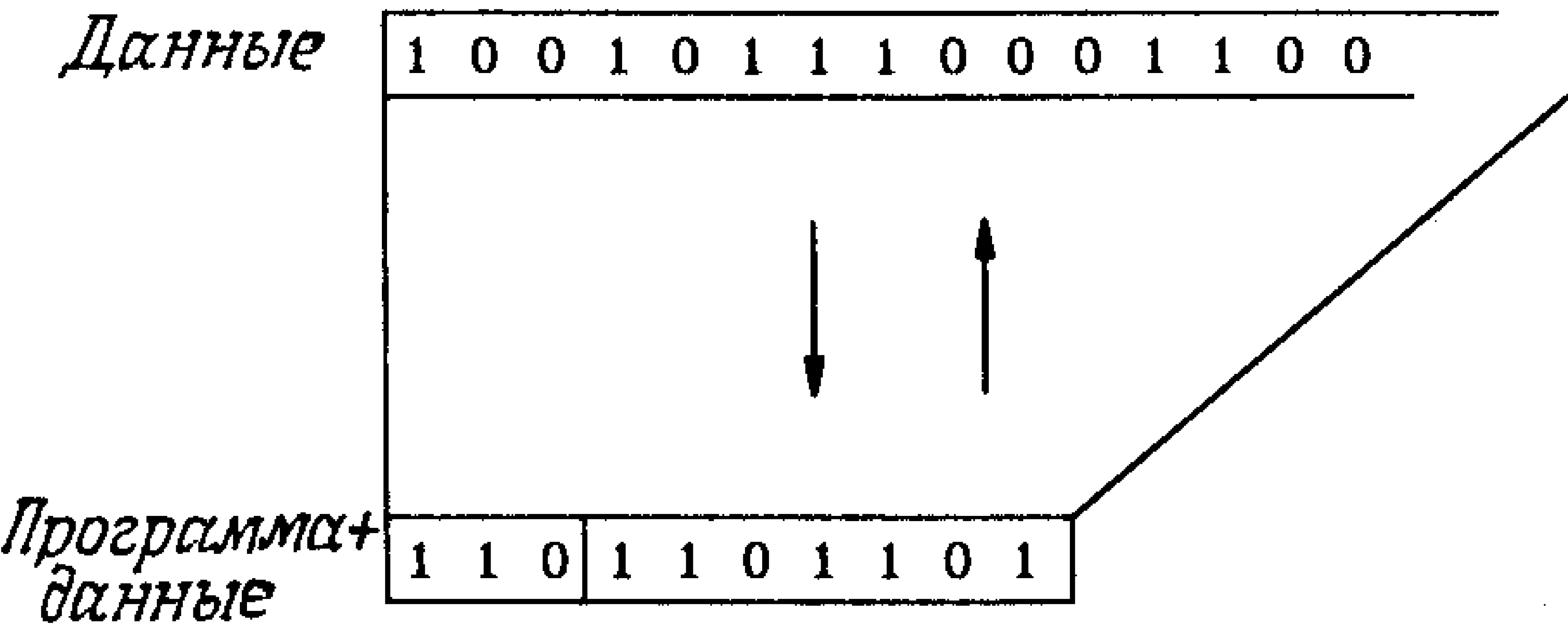


Рис. 1.10. Сжатие строки данных в минимальное множество программы и данных.

априори нужный аспект исследования не представляется возможным, и прежде чем построить более или менее адекватную модель сложной системы, приходится проходить соответствующий курс обучения.

1.2. Как подходить к исследованию сложных систем?

Чем больше дифференцируется наука, распадаясь на отдельные дисциплины, тем большее значение обретает поиск унифицирующих принципов. Так как сложные системы поистине вездесущи, перед нами возникает острая проблема отыскания унифицирующих принципов, которые позволили бы нам должным образом подходить к исследованию таких систем. Для описания сложной системы на микроскопическом уровне необходимо огромное количество данных, которое в настоящее время не в состоянии обработать ни человек, ни даже общество в целом. Следовательно, сбор данных и мышление требуют своего рода экономии. Кроме того, мы надеемся, что, когда нам удастся найти законы, применимые к широкому кругу самых различных сложных систем, на нас снизойдет озарение и мы сможем постичь их внутреннюю сущность.

При поиске универсальных законов разумно спросить, на каком уровне мы хотим их сформулировать – на микроскопическом или на макроскопическом. В зависимости от ответа мы можем прийти к совершенно различному описанию одной и той же системы. Например, на микроскопическом уровне газ совершенно беспорядочен, тогда как на макроскопическом уровне газ практически однороден, т.е. бесструктурен. В отличие от газа кристалл обладает строгой упорядоченностью на микроскопическом уровне, но на макроскопическом уровне также однороден. В биологии мы встречаемся с иерархией уровней от молекулярного уровня через уровни клеток и органов до уровня всего растения или животного. Такое разбиение на уровни может оказаться слишком грубым, и адекватный выбор уровня – задача отнюдь не тривиальная. Кроме того, микроскопичность и макроскопичность уровней становятся относительными понятиями. Например, биомолекулу можно считать макроскопической по сравнению с образующими ее атомами, но микроскопической по сравнению с клеткой. Кроме того, на каждом уровне мы сталкиваемся со специфической организацией или структурой.

Метод современной западной науки может быть охарактеризован как аналитический. Разлагая систему на части, мы пытаемся понять свойства системы как целого. В ряде областей мы можем начать с первых принципов, заложенных в фундаментальных законах. Наиболее отчетливо эта тенденция выражена в таких областях, как физика и особенно физика элементарных частиц. Обычно принято считать, что части и их свойства «объективно» заданы и необходимо «лишь» вывести свойства всей системы как целого из свойств ее частей. В этой связи уместно сделать два замечания. Во-первых,

строго говоря, в действительности мы выводим микроскопические события из макроскопических данных, и было бы интересно проверить, могут ли различные микроскопические модели привести к одному и тому же набору макроскопических данных. Во-вторых, аналитический подход опирается на понятие сводимости или в предельном случае на редукционизм. Но чем больше нам приходится иметь дело с сложными системами, тем яснее мы сознаем, что у редукционизма имеются свои ограничения. Например, из того, что мы знаем химию, отнюдь не следует, будто мы понимаем жизнь. В действительности при переходе с микроскопического уровня на макроскопический перед нами открываются многие новые свойства системы, отсутствовавшие на микроскопическом уровне.

Например, в то время как волну можно описать, указав ее длину и амплитуду, эти понятия чужды такой индивидуальной частице, как атому. Для понимания нам необходимы не поведение отдельных частиц, а их «оркестровка». Чтобы понять оркестровку, мы можем во многих случаях обращаться к модельным системам, на которых специфические особенности сложной системы могут быть подробно исследованы. Ряд таких модельных систем мы обсудим в разд. 1.4. Другой подход к обращению с сложной системой связан с макроскопическим описанием. Например, при описании газа мы не перечисляем координаты каждого из образующих его атомов в отдельности, а используем такие макроскопические величины, как давление и температура. Замечательно, что природа сама наделила нас средствами, позволяющими измерять или ощущать эти величины.

Чтобы иметь дело с сложными системами, нам весьма часто приходится заниматься поиском адекватных переменных или соответствующих величин для описания свойств этих систем. Во всех случаях макроскопическое описание позволяет достигать колоссального сжатия информации, поскольку мы занимаемся рассмотрением не индивидуальных микроскопических данных, а глобальных свойств. Важный шаг в исследованиях сложных систем состоит в установлении соотношений между различными макроскопическими величинами. Эти соотношения являются следствием микроскопических событий, которые, однако, неизвестны или известны только частично. Примеры соотношений такого рода мы находим в термодинамике, где, например, формулируется закон, связывающий давление и температуру в газе, а в статистической механике этот закон выводится из микроскопических законов. В общем случае нам остается лишь догадываться о том, какие микроскопические события приводят в конечном счете к макроскопическим данным.

В этой книге мы намереваемся показать, как можно строить такие догадки относительно систем, изучаемых совершенно различными научными дисциплинами. Одновременно мы увидим, что на достаточно абстрактном уровне между поведением сложных систем существуют глубокие аналогии, или, иначе говоря, что сложное поведение

может быть реализовано на совершенно различных субстратах. Очень часто мы замечаем, что чем сложнее система, тем сильнее сходство между особенностями ее поведения и поведения человека. Это побуждает нас к описанию поведения сложных систем в антропоморфных терминах, хотя такая аналогия может иной раз вводить в заблуждение. В естественных науках стало традицией по возможности изгонять антропоморфизм и основывать все объяснения и понятия на более или менее механистической точке зрения. Мы обсудим дилемму антропоморфного или механистического описания в дальнейших разделах этой главы, в особенности когда мы дойдем до обсуждения информации и роли значения и цели.

Позвольте мне завершить этот раздел книги замечанием общего характера. Не так давно было более или менее широко распространено мнение о том, что между физикой или естественными науками, с одной стороны, и гуманитарными науками, с другой стороны, лежит непреодолимая пропасть и что гуманитарные науки занимаются изучением истинно сложного поведения и сложных систем. Физика долгое время почиталась за свою способность предсказывать события в сколь угодно далеком будущем. Как мы увидим из последующих глав, чем больше физике приходится иметь дело с сложными системами, тем больше мы сознаем необходимость новых понятий. Физика все более утрачивает некоторые черты, считавшиеся ранее ее неотъемлемой принадлежностью, например способность делать точные предсказания.

1.3. Модельные системы

Своими замечательными достижениями физика обязана своей методологии. Согласно принятому в физике подходу, сложную систему подразделяют на те или иные части, поведение которых подвергается воспроизводимому исследованию, причем в ходе изучения части изменяется либо один-единственный параметр, либо весьма небольшое число параметров. Знаменитыми примерами такого подхода могут служить опыты Галилея по свободному падению тел и ньютоновская теория движения планет, построенная путем рассмотрения системы, состоящей только из Солнца и одной планеты. Иначе говоря, Ньютон решал задачу одного тела или самое большее двух тел. Такой подход позволил построить ньютоновскую механику. Некоторые авторы, например Лаплас, опираясь на законы движения Ньютона, пришли к выводу о том, что ньютоновская механика позволяет полностью предсказывать будущее, коль скоро известны скорости и положения всех частиц, образующих систему, в начальный момент времени. В современной физике представление о полной предсказуемости было поколеблено дважды. Квантовая механика утверждает, что скорость и положение частицы невозможно измерить одновременно сколь угодно точно. Следовательно, мы не можем

точно предсказать траекторию частицы в будущем.

В последнее время теория так называемого детерминистического хаоса показала, что даже в классической механике предсказуемость не может быть гарантирована с абсолютной точностью. Рассмотрим следующий очень простой пример: стальной шарик, падающий на вертикально стоящее лезвие бритвы (рис. 1.11). В зависимости от начального положения шарика относительно лезвия его траектория может отклоняться влево или вправо. Это означает, что будущая траектория частицы, т. е. стального шарика, весьма чувствительно зависит от начального условия. Малейшее изменение начального условия может привести к совершенно иной траектории. За последние годы в физике, химии и биологии были обнаружены многочисленные примеры чувствительной зависимости от начальных условий (рис. 1.12). Но, несмотря на эти замечания, общая идея о построении подходящих систем, моделирующих сложную систему, остается в силе.

Мы хотим перечислить лишь несколько наиболее известных примеров модельных систем. Лазер – источник света нового типа – стал

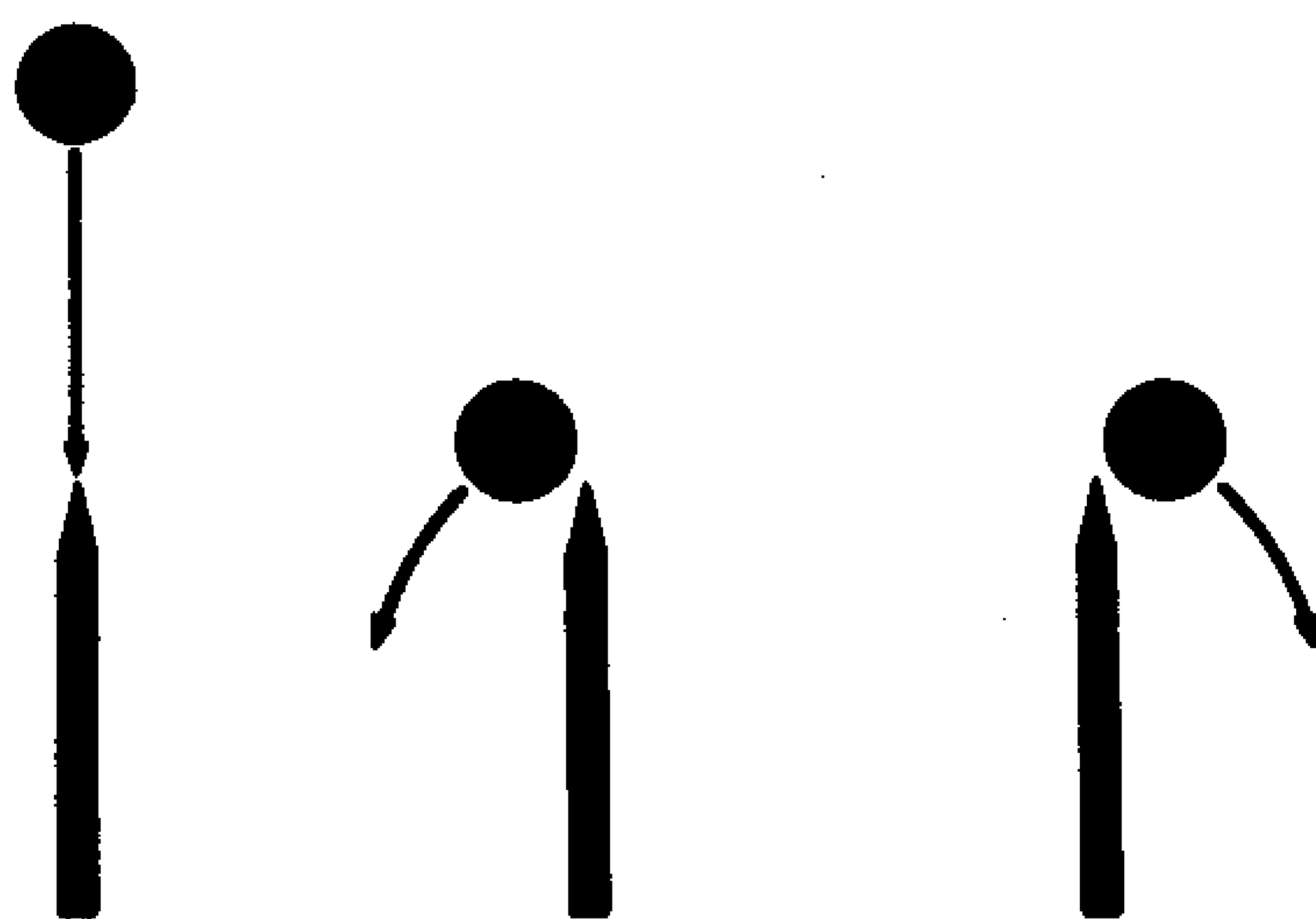


Рис. 1.11. Стальной шарик падает на лезвие бритвы. В зависимости от своего начального положения шарик отклоняется либо влево, либо вправо.

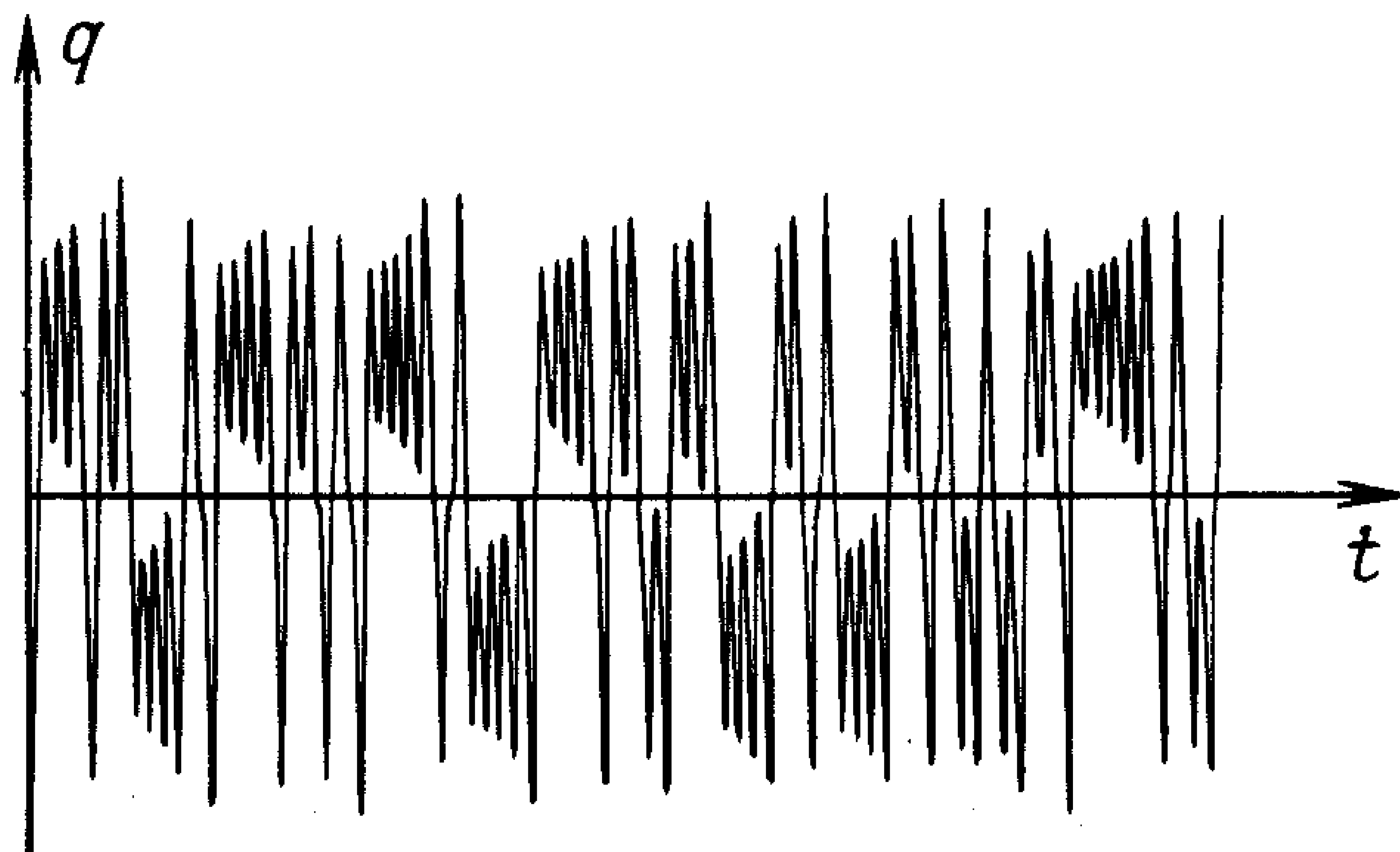


Рис. 1.12. Временная вариация величины в хаотической системе.

парадигмой для самоорганизации когерентных процессов, поскольку в лазере атомы, согласованно взаимодействуя, порождают когерентную лазерную волну (рис. 1.13). Другой пример возникновения макроскопической структуры в результате самоорганизации дает гидродинамика. Например, при подогревании жидкости снизу в ней образуются характерные пространственные структуры, такие, как вихри или ячейки, напоминающие пчелиные соты (рис. 1.14). При более интенсивном нагревании жидкости могут возникнуть пространственно-временные структуры, например осцилляции вихрей. Химические реакции могут также порождать макроскопические структуры – например, химические колебания сопровождаются периодическим изменением цвета с красного на синий и т. д. К числу структур, возникающих в химических системах, относятся также спиральные структуры и концентрические волны (рис. 1.15). В биологии пресноводное животное гидра стала модельной системой для морфогенеза. Если гидру разрезать на две части, то из нижней части быстро формируется новая голова, а из головной части вырастает новая нижняя часть (рис. 1.16).

Тщательные эксперименты могут позволить нам прийти к выводам относительно механизма такой регенерации на основе понятия химических полей, возникающих вследствие производства и диффузии химических веществ. Другим примером самоорганизации в морфогенезе могут служить миксомицеты, или слизневики (рис. 1.17). Эти крохотные грибы обычно существуют в виде отдельных клеток, живущих на субстрате. На протяжении жизненного цикла миксомицета эти отдельные клетки собираются в точку, дифференцируются

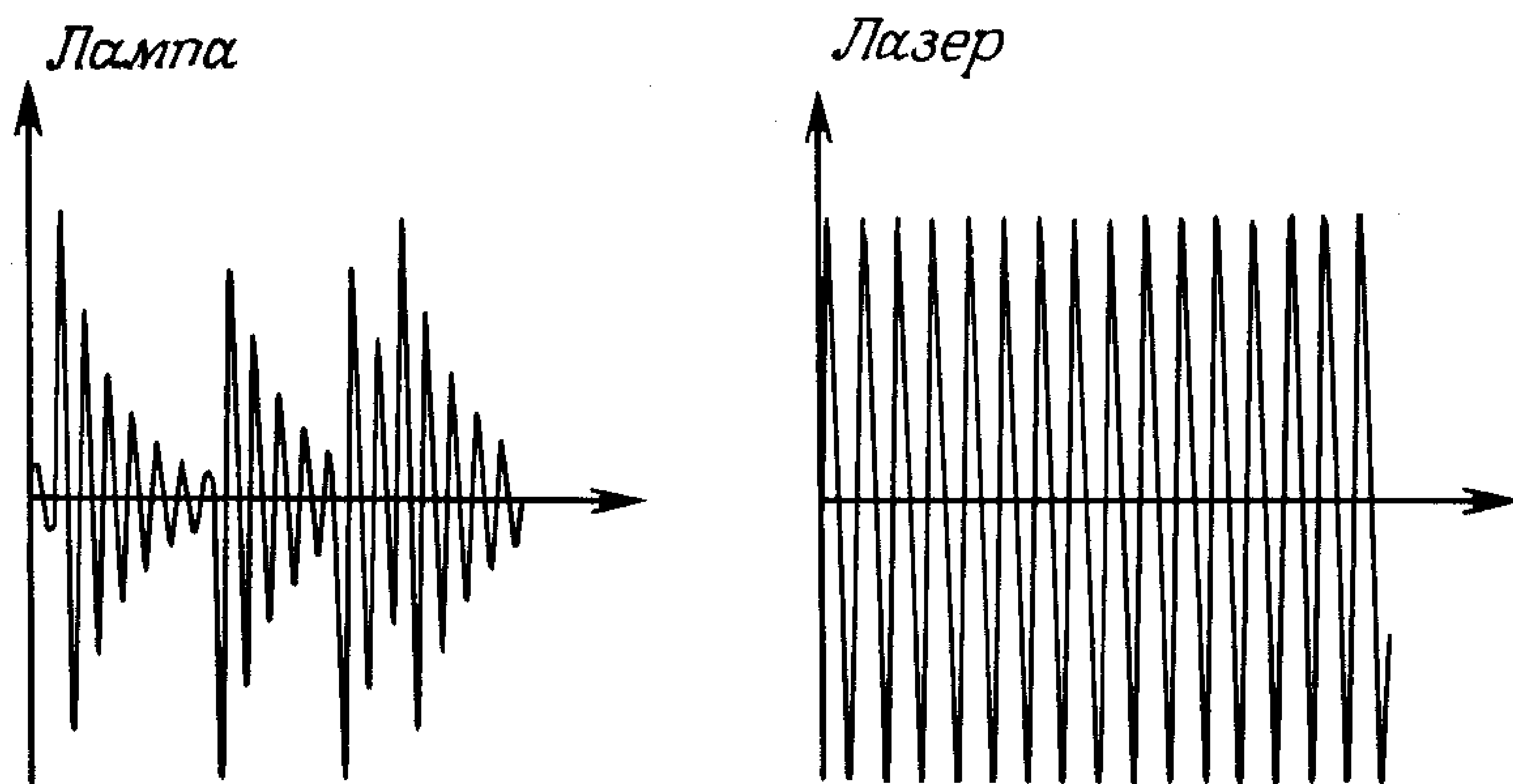


Рис. 1.13. Основное различие между светом, испускаемым лампой, и лазерным излучением. В обоих случаях показана зависимость амплитуды напряженности электрического поля от времени. Испускаемый лампой свет (слева) состоит из отдельных некоррелированных цугов волн. Лазерное излучение (справа) состоит из одного практически бесконечно длинного цуга синусоидальных волн.

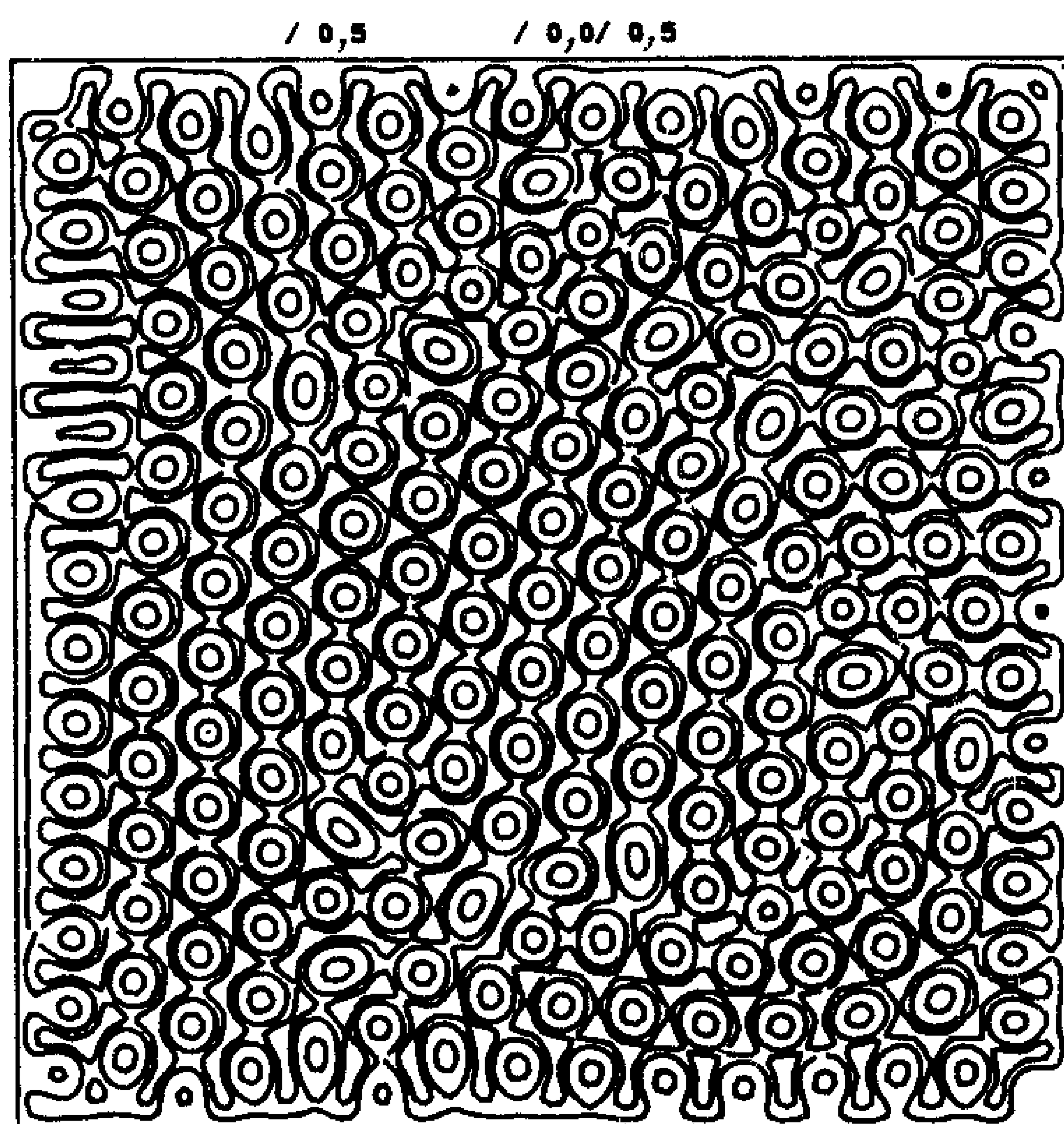


Рис. 1.14. Возникновение структур в гидродинамике.

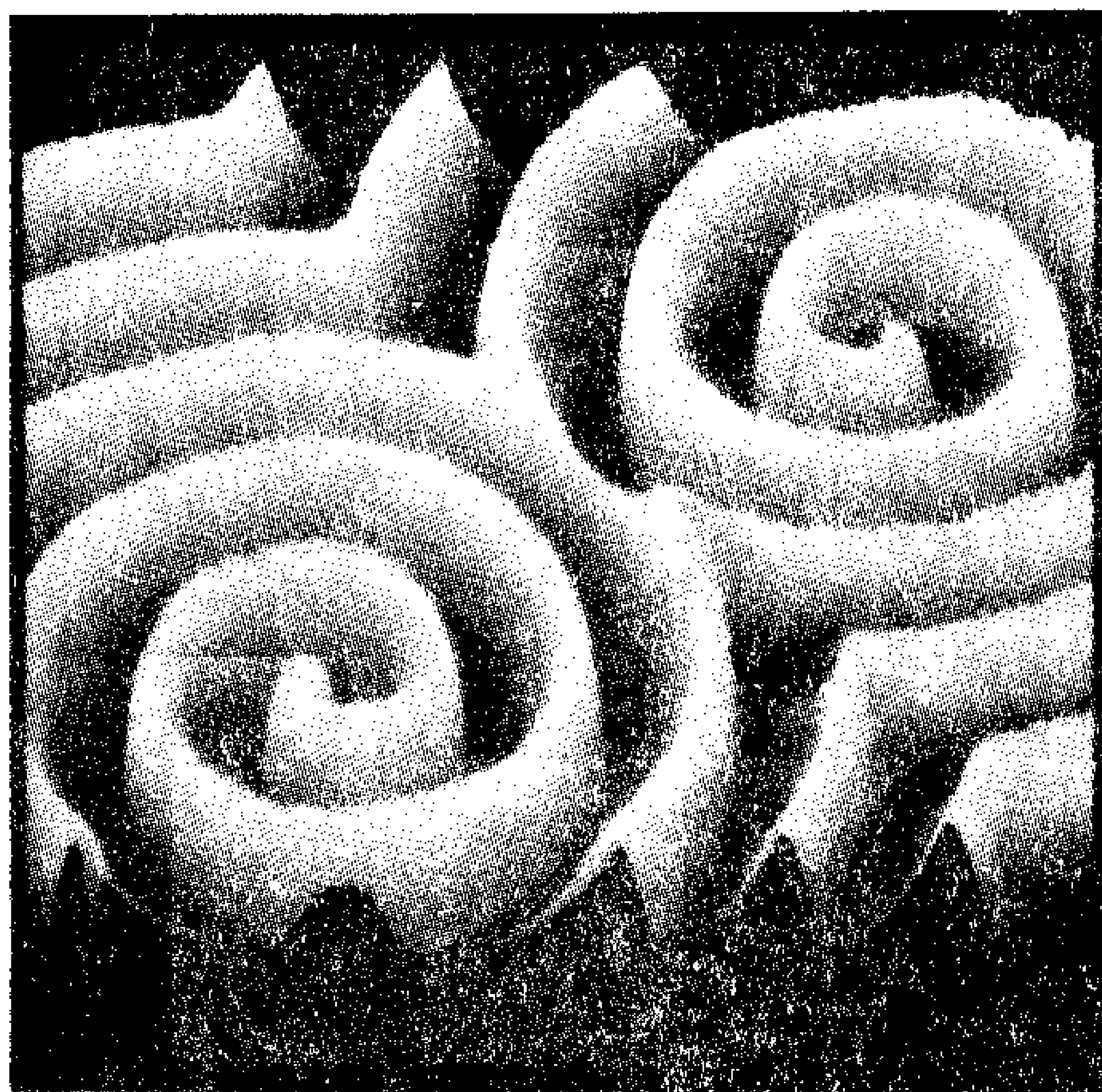


Рис. 1.15. Спиральные волны в реакции Белоусова–Жаботинского. [Из неопубликованной работы С. Ч. Мюллера, Т. Плессера и Б. Хесса.]

и образуют гриб, который затем распространяет споры, после чего жизненный цикл начинается сначала. Еще одна модельная система – аксон кальмара, используемый для изучения нервной проводимости, или хорошо известный пример из генетики: муха дрозофила с ее гигантскими хромосомами, высокой скоростью размножения и легкостью мутаций – идеальный объект исследования.

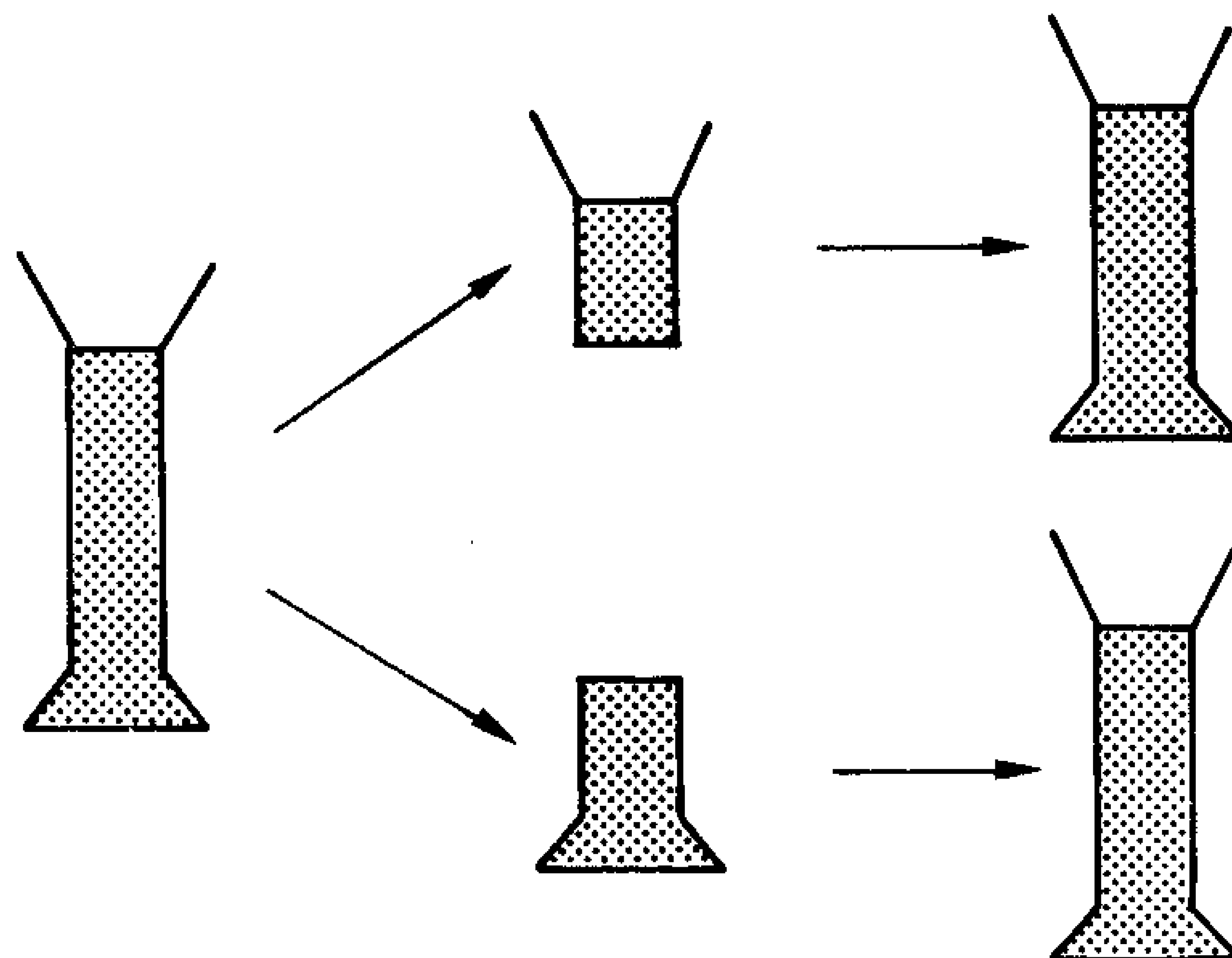


Рис. 1.16. Эксперимент с гидрой показывает, что у этого вида информация о дифференциации клеток не может быть заложена в генах. Слева направо: целая гидра с головной и задней частью перерезается посередине. Через некоторое время верхняя половина регенерирует, образуя недостающую заднюю часть, а нижняя половина регенерирует, образуя недостающую головную часть.

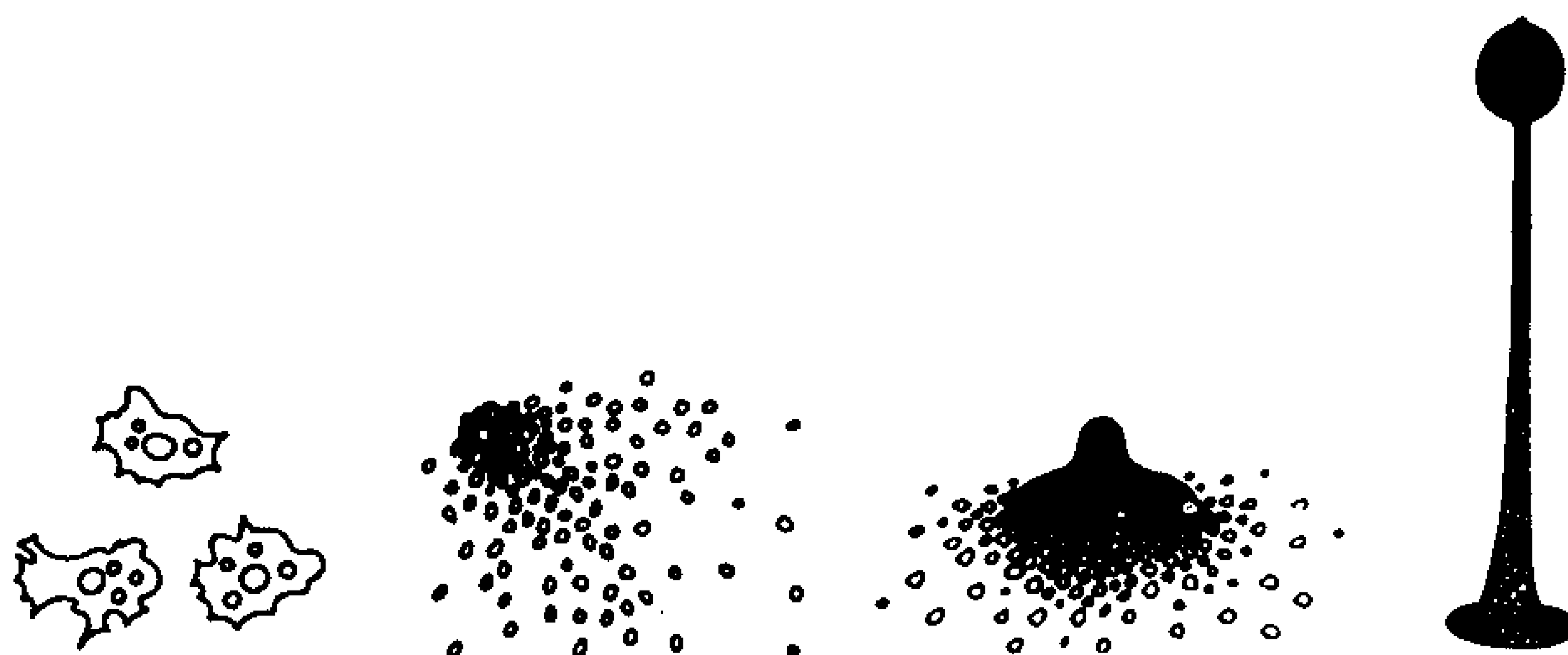


Рис. 1.17. Стадии развития миксомицета (слизневого гриба). Слева направо: отдельные клетки собираются, агрегируют во все более крупное образование и, наконец, дифференцируются. Возникает гриб.

В последнее время движения человеческих рук стали модельной системой для изучения координации мышц и нервных клеток и, в частности, переходов от движений одного типа к движениям другого типа (рис. 1.18). В дальнейшем мы еще вернемся к этому примеру и к другим примерам в этой книге. Непроизвольное изменение характера движений руки сильно напоминает смену аллюров у лошадей, кошек, других четвероногих. В совершенно общем плане такого рода модельные системы позволяют создавать новые понятия, которые сначала подлежат проверке на более или менее широком круге относительно простых систем и лишь затем применяются к подлинно

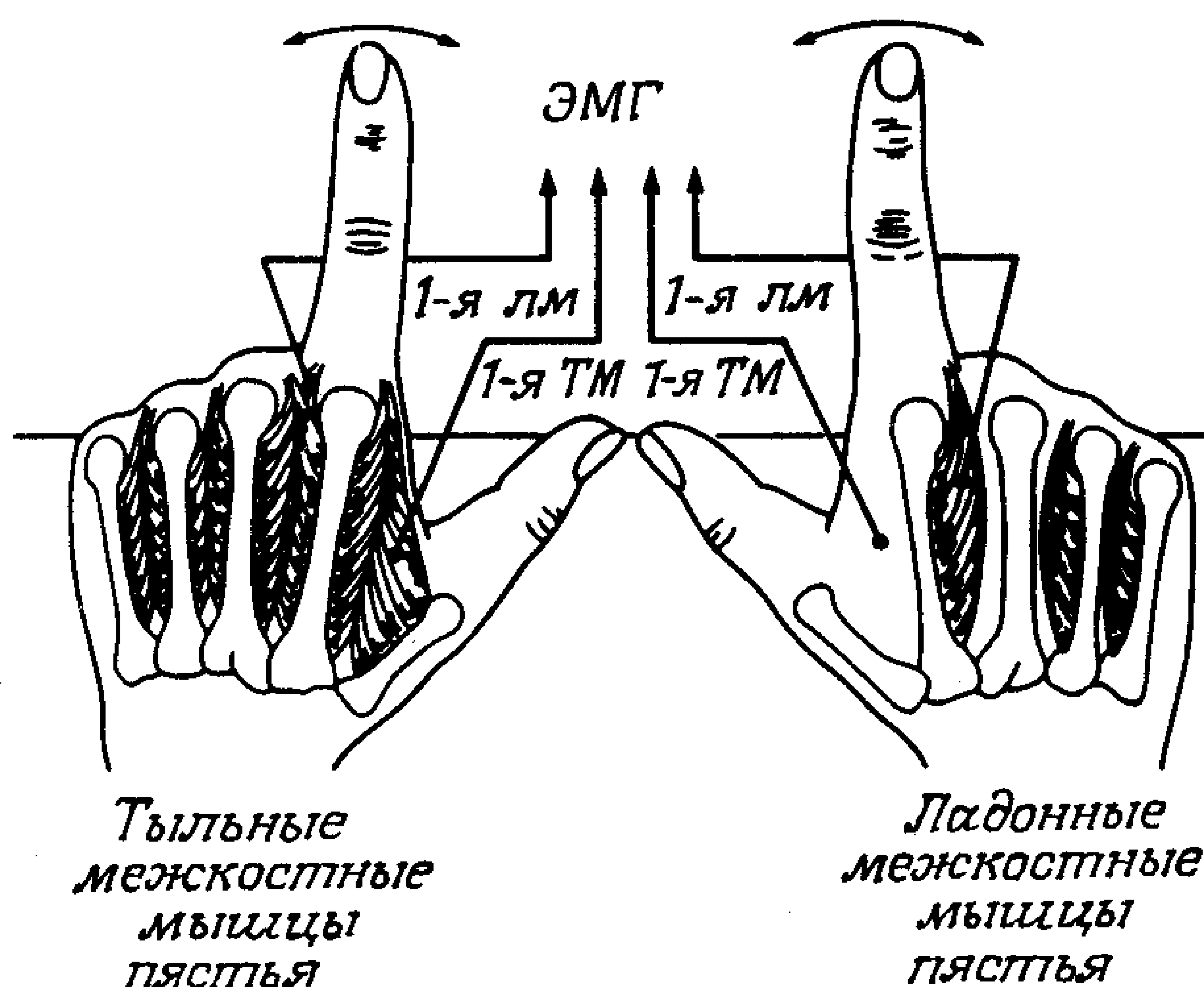


Рис. 1.18. Экспериментальная установка для исследования непроизвольных движений руки (*ЭМГ*–электромиограмма, *ЛМ*–ладонная мышца, *ТМ*–тыльная мышца). [Из работы: *Kelso J. A. S.*, Dynamic Patterns in Complex, Biological Systems: Experiment and Synergetic Theory (препринт).]

сложным системам. Именно в этом смысле последующие главы нашей книги посвящены развитию таких новых понятий, применимость которых будет продемонстрирована на явных примерах.

1.4. Самоорганизация

Как уже упоминалось, существует определенное различие между системами, созданными человеком, и системами, возникшими в результате самоорганизации. В нашей книге мы будем заниматься исследованием самоорганизующихся систем. Следует заметить, однако, что различие между системами, созданными человеком, и самоорганизующимися системами не является четко выраженным. Например, люди могут создавать такие системы, что при наличии определенных ограничений их специфическая функция будет осуществляться путем самоорганизации. Типичным примером такой системы, примером, о котором уже упоминалось выше, может служить лазер. Устройство лазера с его зеркалами позволяет атомам активной среды испускать излучение особого рода. Совершенно очевидно, что когда-нибудь возникнет необходимость в создании компьютеров, самопрограммирующихся на основе самоорганизации.

Для дальнейшего полезно иметь какое-нибудь подходящее определение самоорганизации. Мы называем систему самоорганизующейся, если она без специфического воздействия извне обретает

какую-то пространственную, временную или функциональную структуру. Под специфическим внешним воздействием мы понимаем такое, которое навязывает системе структуру или функционирование. В случае же самоорганизации система испытывает извне неспецифическое воздействие. Например, жидкость, подогреваемая снизу, совершенно равномерно обретает в результате самоорганизации макроструктуру, образуя шестиугольные ячейки. В нашей книге основное внимание будет уделено самоорганизации особого рода, а именно так называемым неравновесным фазовым переходам.

Как известно, системы в состоянии теплового равновесия могут претерпевать определенные переходы из одного состояния в другое при изменении какого-нибудь параметра, например температуры. Так, при нагревании льда он плавится и переходит в новое состояние – жидкость, а именно превращается в воду. При дальнейшем нагревании воды она при определенной температуре закипает и переходит в пар. Таким образом, одни и те же микроскопические элементы, а именно отдельные молекулы, могут образовывать совершенно различные макроскопические состояния, между которыми существуют резкие переходы. При этих переходах возникают качественно новые свойства, например механические свойства льда резко отличаются от механических свойств газа.

В дальнейшем нас будут интересовать аналогичные изменения в состояниях систем, далеких от теплового равновесия. Примеры таких изменений были приведены в разд. 1.3 – это жидкость, образующая пространственную структуру, напоминающую пчелиные соты, лазер, испускающий когерентную световую волну, или биологические ткани, претерпевающие переход к дифференциации, которая приводит к образованию специализированных органов.

1.5. В поисках универсальности

1.5.1. Термодинамика

Термодинамика – одна из областей физики, позволяющая рассматривать произвольно сложные системы с единой точки зрения. Например, мы можем говорить о температуре камня, автомашины, картины или какого-нибудь животного. Мы знаем также, что важные свойства систем изменяются при изменении температуры систем. Вспомним хотя бы таяние льда при температуре плавления или о том, как важно измерять температуру больного. Вместе с тем приведенный нами пример показывает, что одна лишь температура недостаточно полно характеризует автомашину или картину во многих других отношениях. Камень, автомашина, одежда и картина обладают тем свойством, что находятся в тепловом равновесии. Такое состояние достигается, если мы предоставляем систему самой себе или соединяем с другой системой, находящейся в тепловом

равновесии при определенной температуре.

Еще одним важным и, более того, центральным понятием в термодинамике является энтропия. Это понятие относится к системам, находящимся в тепловом равновесии, которое можно охарактеризовать температурой T. Изменение энтропии определяется хорошо известной формулой $dS = dQ_{обр}/T$. Здесь T – абсолютная температура, а $dQ_{обр}$ – количество тепла, обратимо подведенное к системе или отведенное от нее. Общие законы термодинамики принято называть началами. Их два.

1. Первое начало утверждает, что в замкнутой системе энергия сохраняется, причем энергия может принимать различные формы, такие, как внутренняя энергия, совершенная работа или тепло. Таким образом, типичное соотношение, связывающее изменения dU, dA и dQ имеет вид

$$dU = dQ - dA. \tag{1.1}$$

2. Второе начало термодинамики утверждает, что в замкнутой системе энтропия никогда не может убывать, а лишь возрастает до тех пор, пока не достигнет максимума. Как мы увидим в дальнейшем, законы сохранения, например закон сохранения энергии, вместе с так называемым принципом максимума энтропии позволяют нам выводить некоторые микроскопические свойства системы из макроскопических данных. Например, мы можем весьма просто и непосредственно вывести функцию распределения скоростей газа. В этой книге мы будем рассматривать почти исключительно *открытые системы* (рис. 1.19). Существуют системы, которые поддерживаются в определенных состояниях за счет непрерывного притока энергии и/или вещества. Как мы увидим, традиционная термодинамика не позволяет адекватно описывать такие системы; для них нам придется развить некоторую новую разновидность термодинамики, которая будет подробно изложена в последующих главах.

Термодинамику можно рассматривать как макроскопическую феноменологическую теорию. Ее основания лежат в статистической

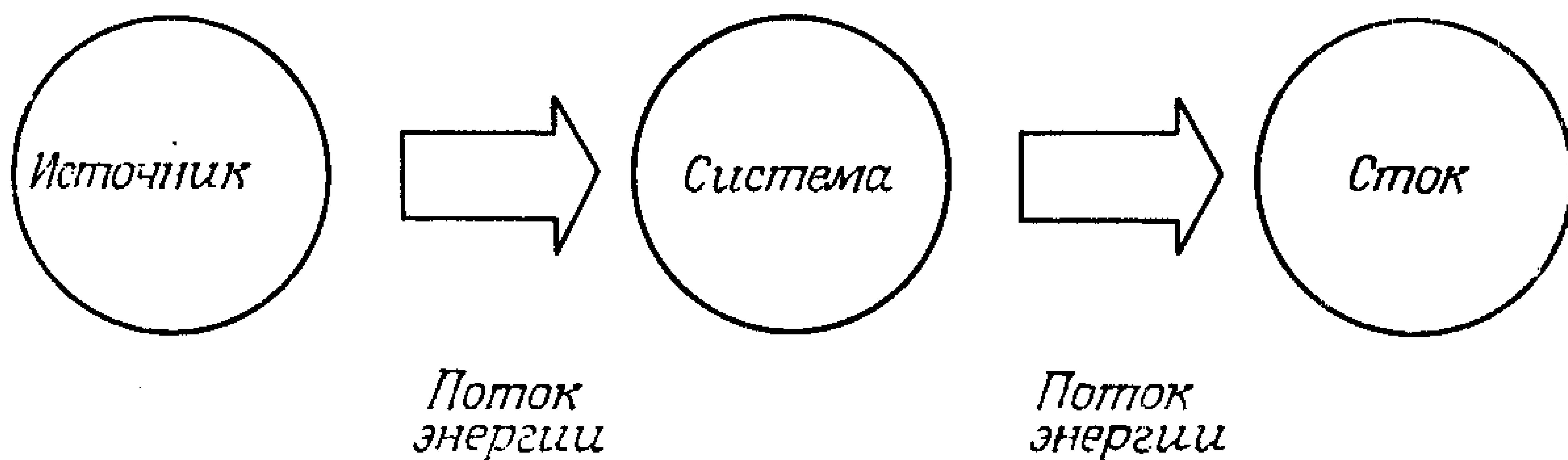


Рис. 1.19. Схема открытой системы, получающей энергию от источника и диссипирующей остаток энергии в сток.

физике, по поводу которой мы сделаем несколько замечаний в следующем разделе.

1.5.2. Статистическая физика

В этой области физики предпринимается, в частности, попытка вывести феноменологические макроскопические законы термодинамики из микроскопической теории. Такой микроскопической теорией может быть ньютоновская механика отдельных частиц газа или квантовая механика. Используя соответствующие статистические средние, мы получаем возможность вывести макроскопические величины из микроскопических законов. Центральным понятием и в этом случае является энтропия S. Согласно Больцману, она связана с числом W различных микроскопических состояний, порождающих *одно и то же макроскопическое* состояние системы соотношением

$$S = k \ln W, \tag{1.2}$$

где k – постоянная Больцмана. Решающее значение имеет так и не получивший убедительного ответа вопрос о том, почему макроскопические явления необратимы, хотя все фундаментальные законы обратимы. Например, законы ньютоновской механики инвариантны относительно обращения времени, т. е. если киноленту пустить в обратную сторону, то все, происходящее на экране в обратной последовательности, не противоречит ньютоновской механике. Вместе с тем совершенно очевидно, что в макроскопической физике процессы необратимые. Например, если у нас есть сосуд с молекулами газа и мы откроем клапан, чтобы газ мог попасть во второй сосуд, то оба сосуда окажутся заполненными газом более или менее равномерно. Однако обратный процесс в природе никогда не наблюдается: никому не доводилось видеть, чтобы второй сосуд самопроизвольно опустел и все молекулы собрались снова в первом сосуде.

Несмотря на трудности, связанные со строгим обоснованием необратимости, статистическая физика позволяет нам объяснить ряд явлений неравновесной термодинамики, такие, как релаксационные процессы, теплопроводность, диффузия молекул и т. д.

1.5.3. Синергетика

Третий подход к формулировке универсальных законов, применимых к сложным системам, – синергетический. В этой области мы изучаем системы, которые могут путем самоорганизации образовывать пространственные, временные или функциональные структуры. В физике синергетика занимается изучением систем, далеких от теплового равновесия. Типичными примерами служат подогреваемая снизу жидкость или лазеры. Химические и биологические системы также допустимо рассматривать как физические системы и применять к ним

синергетический подход. Но синергетика имеет дело и с другими системами, например с такими, изучением которых занимаются экономисты или социологи. В синергетике мы сосредоточиваем внимание на качественных макроскопических изменениях, которые сопровождаются появлением новых структур или функций. Ограничение количественными макроскопическими изменениями – цена, которую приходится платить, чтобы найти общие принципы.

Мы напомним читателю общие принципы синергетики в гл. 2. Затем мы увидим, что в физике синергетика начинается с микроскопической формулировки, например с микроскопических уравнений состояния. В других случаях, в частности в биологии и химии, более уместен мезоскопический подход, при котором мы начинаем с подходящих подсистем, например с адекватных свойств целой клетки в биологии. Предполагается, что на рассматриваемую систему наложены внешние связи, такие, как вполне определенное количество энергии, подводимой к системе. При изменении этого управляющего параметра может возникнуть неустойчивость, и система переходит в новое состояние.

В синергетике показано, что в такой точке потери устойчивости неустойчивыми становится, вообще говоря, небольшое число коллективных мод, которые служат *параметрами порядка*, описывающими макроскопическую структуру. В то же самое время эти макроскопические переменные, т. е. параметры порядка, определяют поведение микроскопических частей системы в силу *принципа подчинения*. Так возникновение параметров порядка и их способность подчинять позволяют системе находить свою структуру. При изменении управляющих параметров в широком диапазоне системы могут проходить через иерархию неустойчивостей и сопровождающих их структур.

Синергетика – область, весьма далекая от завершения, и мы делаем в ней лишь самые первые шаги. За последние 10–20 лет было показано, что поведение многочисленных систем подчиняется общим законам синергетики, и я убежден, что в будущем будет обнаружено много новых тому примеров. С другой стороны, не следует упускать из виду и возможность открытия как новых частных закономерностей, так и еще более общих законов.

Как мы уже упоминали, термодинамика занимается рассмотрением систем, находящихся в тепловом равновесии, тогда как синергетика изучает системы, далекие от теплового равновесия. Но здесь складывается весьма своеобразная ситуация. С одной стороны, мы всегда можем погрузить открытую систему в объемлющую ее замкнутую систему (рис. 1.20). Например, Земля представляет собой открытую систему, так как она получает солнечную энергию и, остывая ночью, отдает энергию в космическое пространство. Но, приняв Солнце и, скажем, часть Вселенной за целую систему, мы можем рассматривать эту систему как замкнутую, к которой при-



Рис. 1.20. Открытая система, погруженная в замкнутую систему.

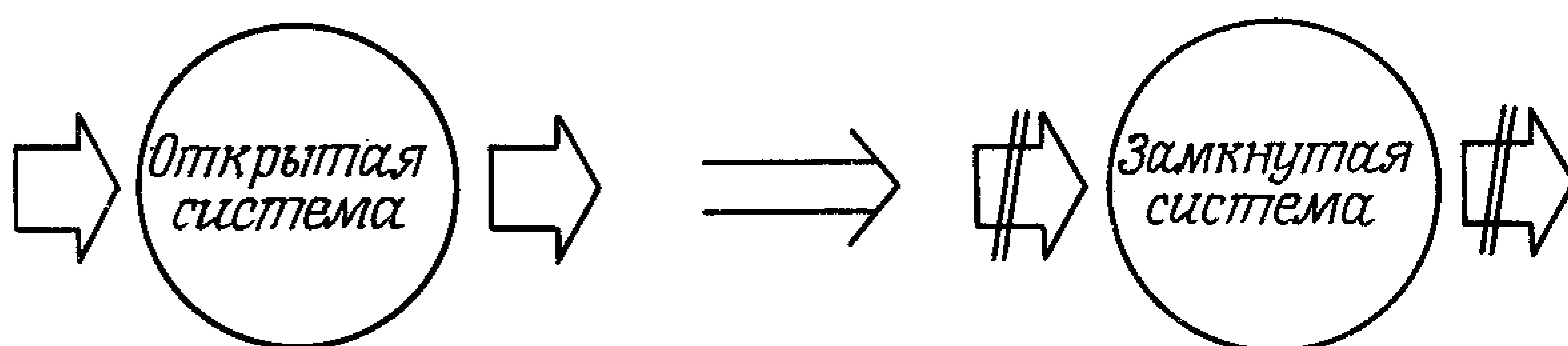


Рис. 1.21. Замкнутую систему можно рассматривать как предельный случай открытой системы, у которой пресечены потоки энергии на входе и выходе.

менимы законы термодинамики. Но, с другой стороны, любую открытую систему можно рассматривать в пределе, когда потоки энергии или вещества стремятся к нулю, и мы в конечном счете имеем дело с замкнутой системой (рис. 1.21). Следовательно, общие законы термодинамики должны получаться как предельные случаи из общих законов синергетики.

Как читатель, должно быть, заметил, эта программа отнюдь не завершена и оставляет место для будущих исследований. До сих пор синергетика начинала свои построения с микроскопического или мезоскопического уровня. Но в этой книге мы попытаемся изложить другие основания синергетики; они сначала частично обсуждены в гл. 5–7, а затем более подробно в последующих главах. Исходным пунктом для этого макроскопического подхода является понятие информации, и мы рассмотрим некоторые из его наиболее важных аспектов в следующем разделе.

1.6. Информация

Использование слова «информация» приводит к многим недоразумениям. Это связано с тем, что оно имеет много различных значений. В

обыденном языке это слово используется в смысле «сообщение» или «сведения». Письмо, телевизионная передача или телефонный разговор несут информацию. В дальнейшем мы будем использовать слово «информация» в его научном значении. Мы начнем с понятия шенноновской информации, согласно которому информация оценивается независимо от ее смысла. Затем мы кратко рассмотрим информацию в связи с передачей сигналов и, наконец, коснемся проблемы самозарождения смысла.

Совершенно очевидно, что при рассмотрении физических систем мы стремимся исключить всякого рода антропоморфизм, поскольку хотим описывать физическую систему как можно более объективно. Но и в биологии эта тенденция столь очевидна, что в конце концов мы получаем более или менее физическую или даже механистическую картину биологических систем. Однако, когда речь заходит о развитии современных компьютеров, например компьютеров пятого поколения, мы стремимся ввести в эту область столь антропоморфные понятия, как смысл, значимость и т. д. Поэтому в этом разделе мы обсудим те направления, следуя которым можно вернуться от понятия информации, лишенной смысловой нагрузки, к акту самозарождения смысла.

1.6.1. Шенноновская информация: изгнание смыслового содержания

Подробно мы рассмотрим понятие информации по Шеннону в гл. 3, но для того, чтобы иметь прочную основу для последующего обсуждения, поясним пока это понятие на нескольких примерах. Когда мы бросаем монету, возможны два исхода. Когда мы бросаем игральную кость, возможны шесть исходов. При бросании монеты мы получаем информацию двух типов: орел или решка. При бросании игральной кости мы получаем информацию о выпадении одного из шести чисел от 1 до 6. Аналогичным образом мы можем, задавая вопросы, получать ответы типа «да» или «нет» и т. д. Понятие информации по Шеннону относится просто к числу возможных исходов, или ответов. Для бросания монеты $Z = 2$, для бросания игральной кости $Z = 6$. Как мы увидим в дальнейшем, собственно мерой информации служит не само число Z, а его логарифм, причем обычно принято брать логарифм по основанию 2, т. е. информация по определению равна величине

$$I = \log_2 Z. \tag{1.3}$$

Этому определению можно придать и другую форму, с которой мы будем время от времени встречаться в этой книге. Рассмотрим, например, такой язык, как английский. Буквы его алфавита $a, b, c, \ldots$ можно обозначить числами $j = 1, 2, 3, \ldots$, т. е. установить соответствие $a \to 1$, $b \to 2$ и т. д. Затем мы можем подсчитать частоты N_j, с

которыми встречаются в какой-нибудь книге или, быть может, библиотеке буквы a, b, c, ... Определим относительную частоту буквы, обозначенной нами j, как величину

$$p_j = \frac{N_j}{N}, \tag{1.4}$$

где N – общее число подсчитанных букв: $N = \sum N_j$. Средняя информация, приходящаяся на одну букву в обследованной книге (или библиотеке) определяется выражением

$$i = -\sum_j p_j \log_2 p_j. \tag{1.5}$$

Вывод этой формулы мы приводим в гл. 3. Шеннон использовал определение информации (1.5) при изучении пропускной способности канала связи – способности передавать информацию даже при наличии помех. Для дальнейшего особое значение имеют две особенности шенноновской информации. 1) Шенноновская информация никак не связана со смыслом передаваемого сигнала. В его концепцию информации не входят такие ее аспекты, как осмысленность или бессмысленность, полезность или бесполезность и т. д. 2) Шенноновская информация относится к замкнутым системам. Имеется ограниченный резервуар сигналов, число которых равно Z.

1.6.2. Как информация действует на систему

В этом разделе мы хотим кратко изложить новый подход, который можно рассматривать как шаг к концепции информации, включающей семантику. К основной идее этого подхода приводит следующее замечание: смысл сигналу можно приписать только в том случае, если мы примем во внимание отклик того, кто принял сигнал. Так мы приходим к понятию *относительной значимости* сигналов, которое мы хотим изложить ниже.

Рассмотрим множество сигналов, каждый из которых закодирован строкой чисел. Основная проблема состоит в моделировании приемника. Мы решаем ее, привлекая современные понятия теории динамических систем, или, более общо, понятия синергетики. Математическое описание таких систем мы дадим в следующей главе, а пока для наших целей достаточно нескольких общих замечаний. Рассмотрим какую-нибудь систему, например газ, биологическую клетку или экономику, состояния которой могут быть охарактеризованы на микроскопическом, мезоскопическом или макроскопическом уровне набором величин q, которые мы перенумеруем индексом j, т. е. величин q_j. Разумеется, величины q_j могут изменяться со временем. Все q_j мы можем соединить в вектор состояния $\mathbf{q}(t) = [q_1(t), q_2(t), ...$

..., $q_N(t)$]. Эволюция вектора $\mathbf{q}$ со временем, т.е. динамика системы, определяется дифференциальными уравнениями вида

$$\frac{d\mathbf{q}}{dt} = \mathbf{N}(\mathbf{q}, \alpha) + \mathbf{F}(t), \tag{1.6}$$

где $\mathbf{N}$ – детерминистическая часть, а $\mathbf{F}$ – флуктуирующие силы. Пока все, что нам необходимо знать, сводится к следующему. Если в отсутствие флуктуирующих сил значение вектора состояния $\mathbf{q}$ в начальный момент времени известно и так называемые управляющие параметры α заданы, то будущее вектора $\mathbf{q}$ определено однозначно. Со временем вектор $\mathbf{q}$ стремится выйти на аттрактор. Чтобы наглядно представить себе простой пример такого аттрактора, рассмотрим миниатюрный рельеф с возвышенностями и впадинами, выполненный из листа бумаги (рис. 1.22, 1.23). Задание управляющих параметров α означает определенный выбор рельефа, по которому шарик может скатываться под действием силы тяжести (и силы трения). Задание вектора состояния $\mathbf{q}$ в начальный момент времени означает, что шарик в начальный момент времени помещают в некоторую вполне определенную точку, например на склоне холма (рис. 1.22). Он скатывается оттуда все ниже и ниже, пока не достигает самой глубокой точки впадины, которая и является аттрактором в этом случае. Специалистам известно, что у динамических систем могут быть аттракторы и других типов, например предельные циклы, когда система совершает незатухающие колебания, или еще более сложные аттракторы, известные под названием *хаотических аттракторов*. В дальнейшем вполне достаточно наглядно представлять себе аттрактор как самую глубокую точку впадины (так называемую неподвижную точку). При наличии флуктуаций $\mathbf{F}$ шарик может перепрыгивать из одного аттрактора в другой (рис. 1.24).

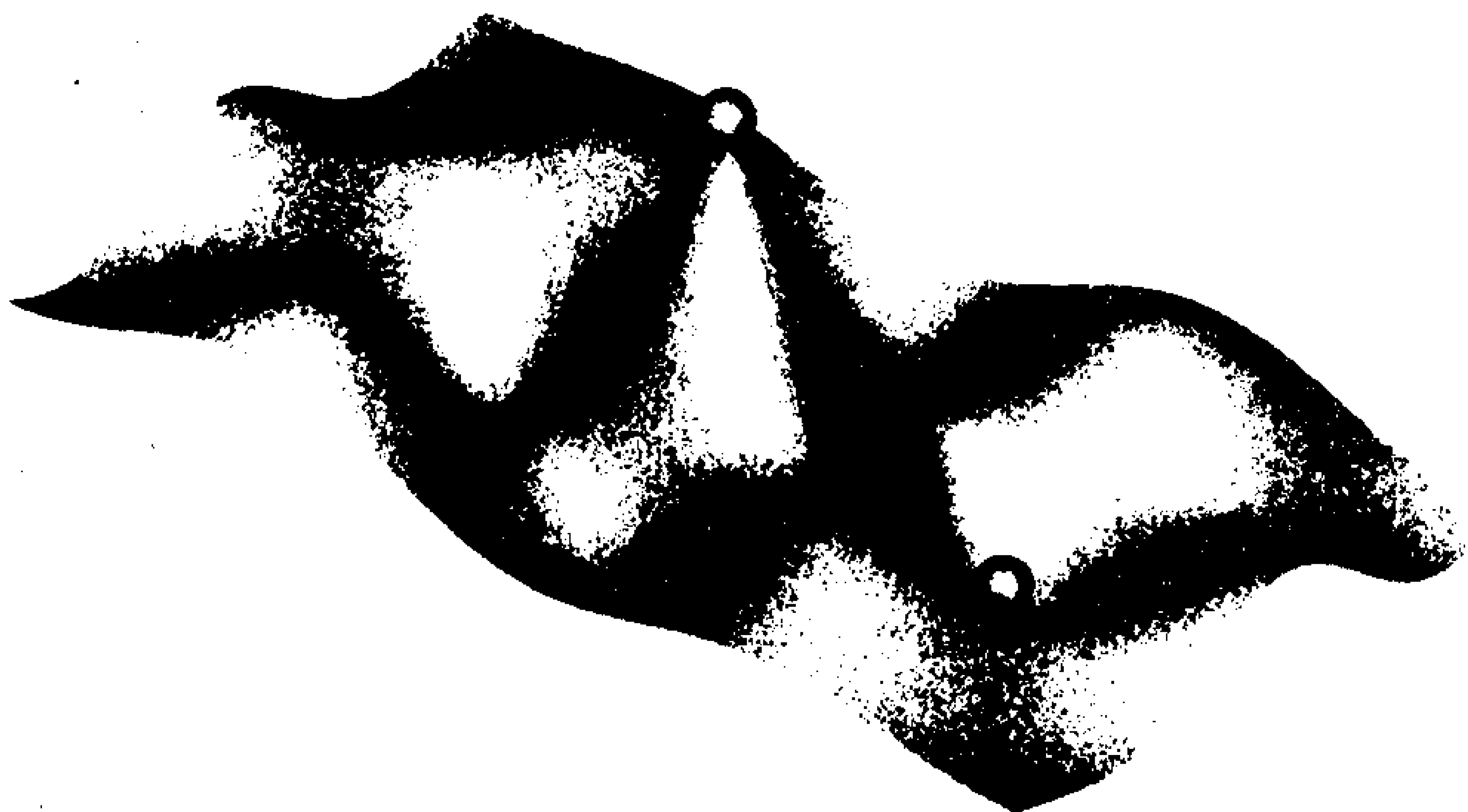


Рис. 1.22. Наглядная модель динамической системы с притягивающими неподвижными точками (аттракторами) – миниатюрный рельеф, выполненный из надлежащим образом изогнутого листа бумаги.

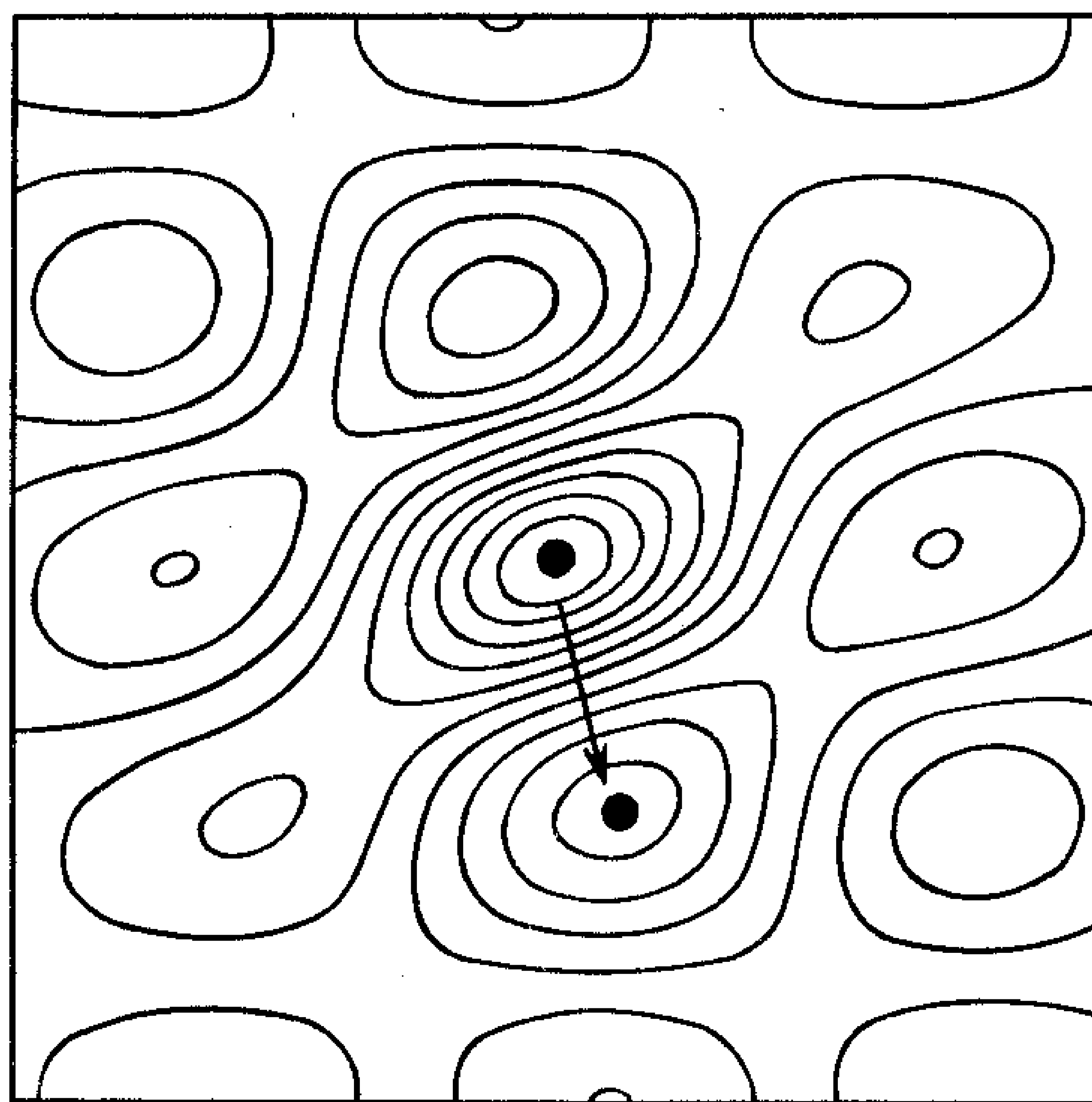


Рис. 1.23. Горизонтали рельефа, изображенного на рис. 1.22.

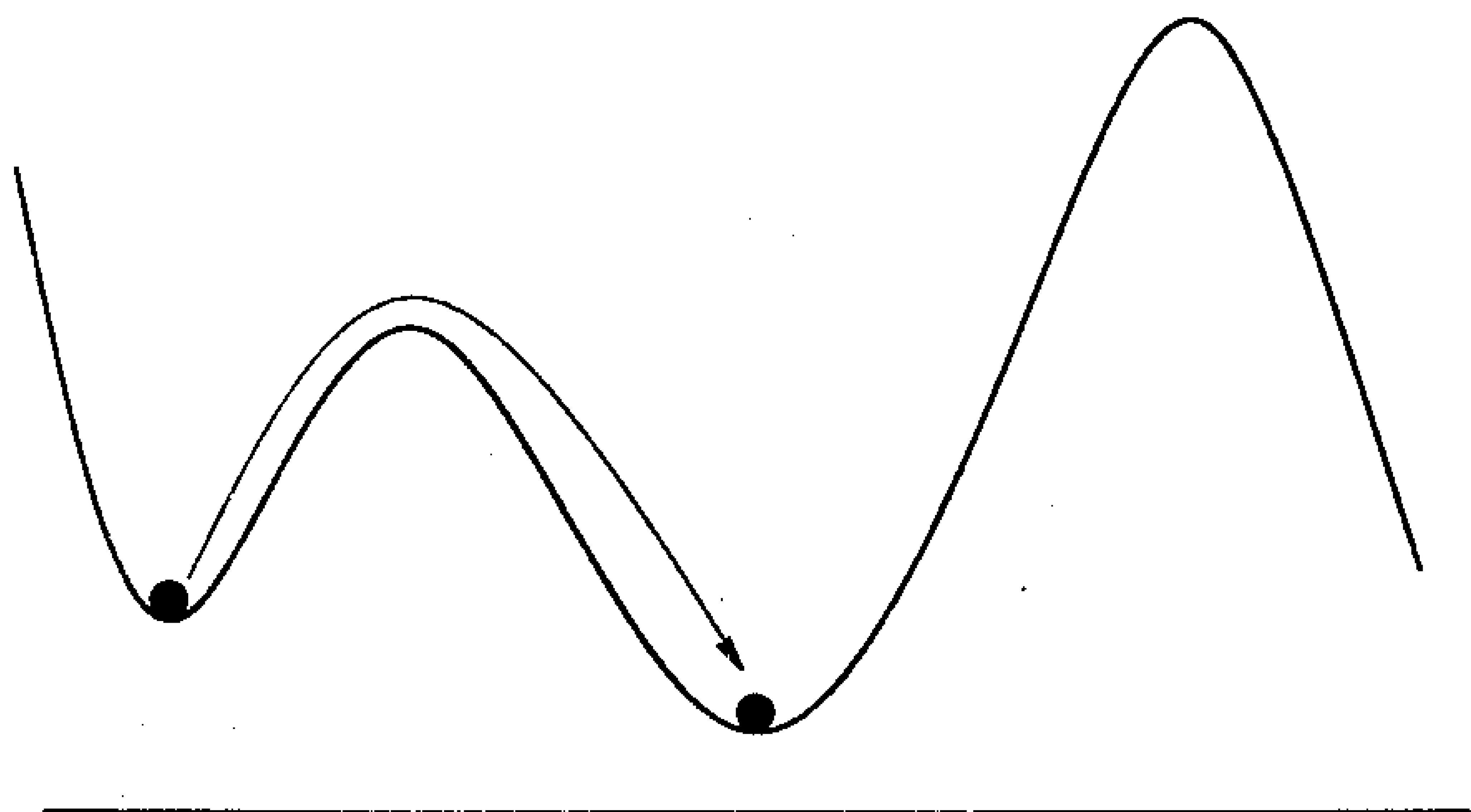


Рис. 1.24. Переход системы от одного аттрактора (притягивающей неподвижной точки) к другому.

После этих предварительных замечаний вернемся к нашей первоначальной проблеме, а именно к проблеме придания смысла сигналу.

Мы предполагаем, что получение сигнала системой означает задание с помощью сигнала управляющих параметров α и начального значения вектора $\mathbf{q}$. Условимся пока считать, что тем самым α и $\mathbf{q}$ однозначно определены. Развиваемая далее теория допускает непосредственное обобщение на случай неполного сигнала (см. ниже). Будем сначала игнорировать роль флуктуаций. Мы предполагаем, что до получения сигнала система находилась в аттракторе, т.е. пребывала в состоянии, которое мы будем называть нейтральным состоянием. Аттрактор может быть и состоянием покоя, т.е. не-

подвижной точкой, и предельным циклом, тором или странным аттрактором или даже аттрактором пока неизвестного типа, который еще предстоит открыть теории динамических систем. Обозначим аттрактор, в котором система пребывает до получения сигнала через $\mathbf{q}_0$. После того как сигнал получен и параметры α и начальное значение вектора $\mathbf{q}_0$ установлены, в принципе возможны два случая. Предположим, что мы можем выжидать в течение некоторого времени измерения, пока динамическая система не перейдет в один из своих возможных аттракторов. Тогда один из двух возможных случаев состоит в том, что полученный сигнал оставляет систему в исходном состоянии $\mathbf{q}_0$. В этом случае сигнал, очевидно, бесполезен или бессмыслен.

В другом случае система переходит в новый аттрактор. Предположим сначала, что этот аттрактор однозначно определяется поступающим сигналом. Ясно, что различные сигналы могут приводить систему в один и тот же аттрактор. Мы говорим тогда об избыточности сигналов.

Наконец, вплоть до недавнего времени считалось загадкой, особенно в области биологии, каким образом может возникать информация. Ответ на этот вопрос нетрудно представить себе наглядно, если предположить, что приходящий сигнал порождает ситуацию, изображенную на рис. 1.25 и заведомо неоднозначную. В зависимости от флуктуаций внутри самой системы последняя может переходить в одну из двух устойчивых точек (в один из двух аттракторов). В этом случае приходящий сигнал содержит информацию в обычном смысле слова, который неоднозначен, и эта неоднозначность снимается флуктуацией системы. Если говорить нестрого, то можно сказать, что исходная информация удваивается, так как после получения сигнала система может перейти в любой из двух аттракторов. В физике мы в подобных случаях говорим об эффектах нарушения симметрии.

Взятые вместе все эти различные процессы позволяют нам перечислить возможные элементарные схемы (рис. 1.26). Разумеется, если мы хотим учесть воздействие других сигналов, то нам могут понадобиться более сложные схемы, например такие, как на рис. 1.27.

Перейдем теперь к вопросу о том, каким образом мы можем

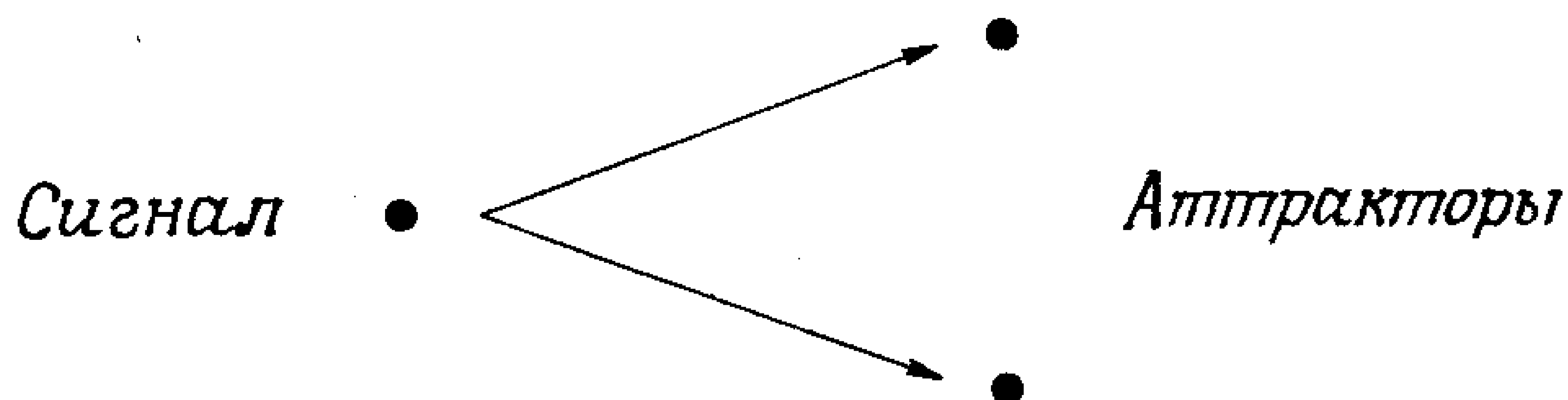


Рис. 1.25. Сигнал может достигать одного из двух различных аттракторов под действием внутренних флуктуаций системы в результате процесса, изображенного на рис. 1.24. При этом становятся доступными два аттрактора.

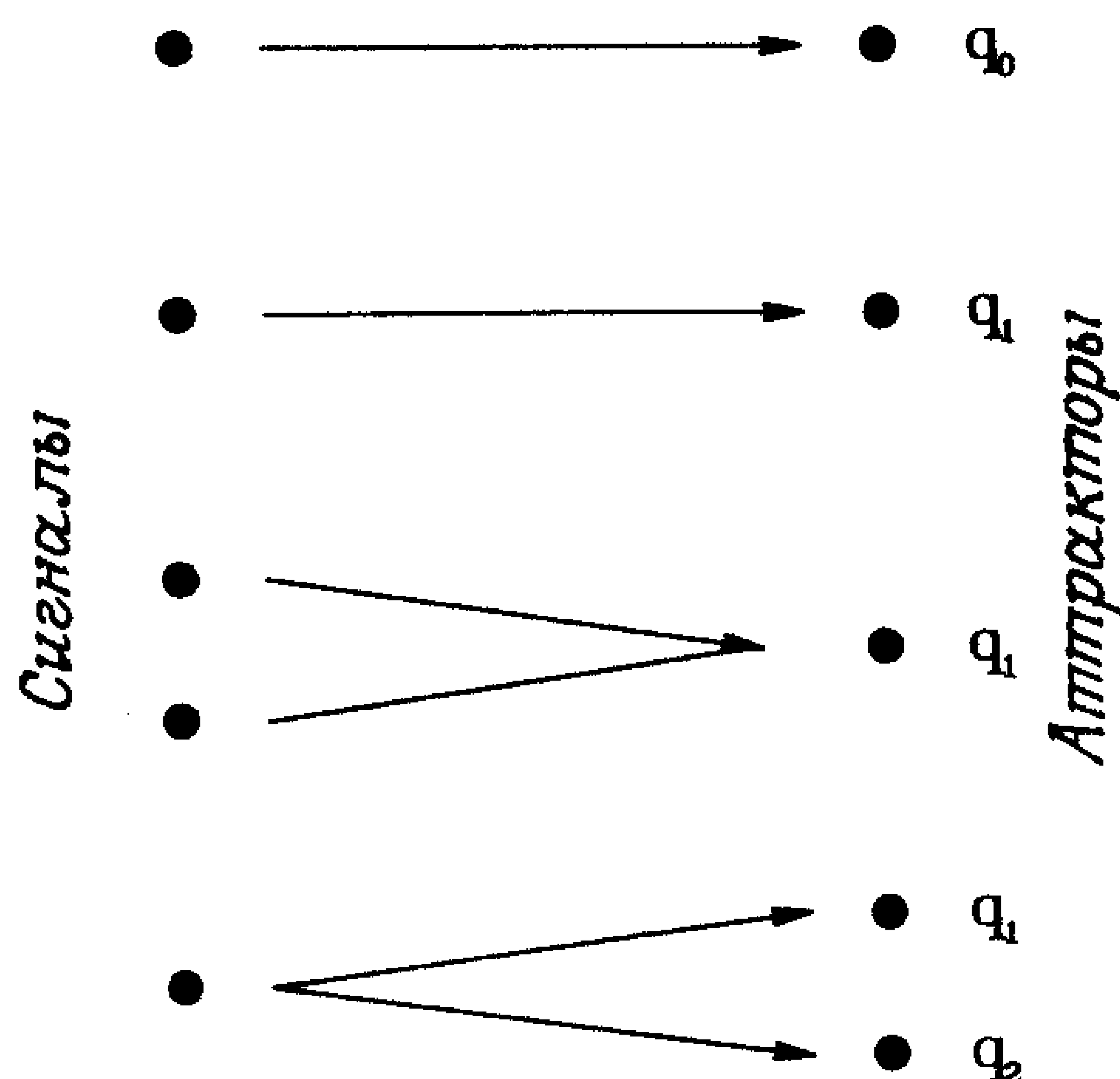


Рис. 1.26. Различные пути, по которым сигнал может достичь аттракторов.

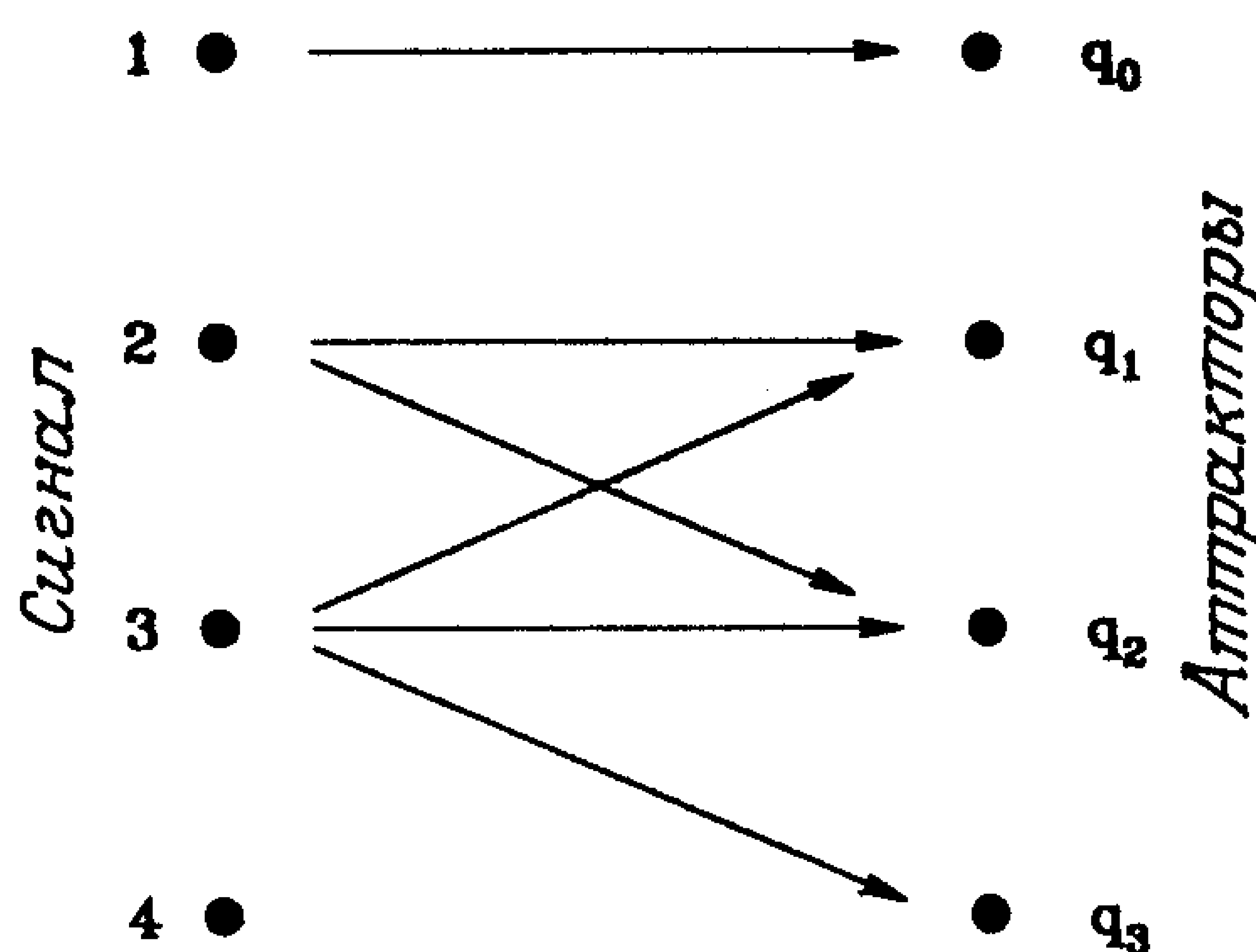


Рис. 1.27. Еще один путь, по которому сигналы могут достичь аттракторов.

приписать значения приходящим сигналам, или, точнее говоря, определить *относительную значимость сигналов*. Для этого мы сначала введем *относительную значимость* отдельных аттракторов. В действительности отдельные аттракторы становятся источниками новых сигналов, которые поступают в новую динамическую систему, и этот процесс может продолжаться до бесконечности. Но для практических целей эту иерархическую последовательность приходится обрывать на некотором уровне, и на этом уровне мы должны приписать отдельным аттракторам определенные значения – степени их значимости. Так как наша процедура становится совершенно ясной даже в случае одноступенчатого процесса, рассмотрим такой процесс подробно.

Припишем *относительную значимость* отдельным аттракторам так, чтобы аттрактор 0 с $\mathbf{q}_0$ имел относительную значимость, равную 0, а остальные аттракторы имели значимость $0 \leqslant p'_i \leqslant 1$, где величины удовлетворяют условию нормировки

$$\sum_j p'_j = 1. \tag{1.7}$$

Выбор величин p'_j зависит от того, какая задача стоит перед динамической системой. В одном случае речь может идти о какой-то конкретной задаче, которая может быть выполнена всего лишь одним аттрактором, в другом – об ансамбле задач, выполнение которых имеет заданную относительную значимость. Ясно, что относительная значимость p_j сигналов зависит не только от динамической системы, но и от тех задач, которые ей предстоит решить. Возникает вопрос: каковы значения p_j приходящих сигналов? Чтобы ответить на этот вопрос, рассмотрим связи между сигналом и аттрактором, в который вынуждена перейти динамическая система по получении данного сигнала. Если система по получении сигнала j переходит в аттрактор k (в частности, в аттрактор 0), то мы ставим в соответствие этому процессу матричный элемент $M_{jk} = 1$ (в противном случае $M_{jk} = 0$). Если в системе возможны внутренние флуктуации, то один-единственный сигнал с примесью флуктуаций может перевести систему в несколько различных аттракторов, выбор которых определяется коэффициентами ветвления M_{jk}, удовлетворяющими условию $\sum_k M_{jk} = 1$. Относительную значимость p_j сигнала мы определим соотношением

$$p_j = \sum_k L_{jk} p'_k = \sum_k \frac{M_{jk}}{\sum_{j'} M_{j'k} + \varepsilon} p'_k \tag{1.8}$$

при условии $\varepsilon \to 0$. (Это условие необходимо для того, чтобы отношение, стоящее в правой части формулы (1.8), имело смысл, даже если числитель и знаменатель одновременно обращаются в нуль.) Предположим сначала, что при любом $p'_k \neq 0$ по крайней мере один коэффициент ветвления $M_{jk} \neq 0$. Нетрудно убедиться в том, что относительные значимости p_j нормированы. Показать это можно по разделениям:

$$\sum_j p_j = \sum_{k,j} \frac{M_{jk}}{\sum_{j'} M_{j'k} + \varepsilon} p'_k = \tag{1.9}$$

$$= \sum_k \left(\sum_j \frac{M_{jk}}{\sum_{j'} M_{j'k} + \varepsilon} \right) p'_k = \tag{1.10}$$

$$= \sum_k p'_k = 1, \tag{1.11}$$

где сумма, стоящая в равенстве (1.10) в скобках, равна 1.

Рассмотрим теперь случай, когда при каких-то значениях k, для которых $p'_k \neq 0$, все $M_{jk} = 0$. В этом случае суммы по k в (1.9) и (1.10) некоторые коэффициенты при $p'_k \neq 0$ равны нулю, а так как $\sum_k p'_k = 1$, мы приходим к неравенству $\sum_j p_j < 1$. Если это неравенство выполняется, мы говорим о *недостаточности*, или неполноте, *информации*.

При более абстрактном подходе мы можем принять левую часть соотношения (1.8) за основу определения, предположив, что

$$\sum_j L_{jk} \leqslant 1, \tag{1.12}$$

где равенство имеет место только в том случае, если нет недостаточности информации.

Заметим, что вместо требования $M_{jk} = 1$, как в случае одного конечного аттрактора для приходящего сигнала, можно считать, что коэффициенты M_{jk} удовлетворяют обобщенному требованию

$$0 < M_{jk} \leqslant 1. \tag{1.13}$$

Формула левой части соотношения (1.8) позволяет нам сразу же выписать формулы для нескольких систем, соединенных последовательно. Например, для двухступенчатого процесса мы получаем без всяких затруднений

$$p_j = \sum_k L^{(1)}_{jk} p'_k = \sum_{kk'} L^{(1)}_{jk} L^{(2)}_{kk'} p''_{k'}. \tag{1.14}$$

Нетрудно убедиться, что $\sum_j p_j = 1$, если $\sum_k p'_k = 1$ и $\sum_j L_{jk} = 1$. Отдельные этапы вычислений выглядят так:

$$\sum_j p_j = \sum_j \sum_{kk'} L^{(1)}_{jk} L^{(2)}_{kk'} p''_{k'} = \tag{1.15}$$

$$= \sum_{kk'} \underbrace{\left(\sum_j L^{(1)}_{jk} \right)}_{=1} L^{(2)}_{kk'} p''_{k'} =$$

$$= \sum_{k'} \underbrace{\sum_k L^{(2)}_{kk'}}_{=1} p''_{k'} = 1. \tag{1.16}$$

Мы можем положить по определению

$$L'_{jk'} = \sum_k L^{(1)}_{jk} L^{(2)}_{kk'}. \tag{1.17}$$

Так как величины L положительны, справедливо неравенство

$$L'_{jk} \geqslant 0, \tag{1.18}$$

а вследствие свойств нормировки (в отсутствие недостаточности информации)

$$\sum_j L'_{jk'} = \sum_k \sum_j L^{(1)}_{jk} L^{(2)}_{kk'} =$$

$$= \sum_k L^{(2)}_{kk'} = 1, \tag{1.19}$$

откуда мы получаем, что

$$L'_{jk} \leqslant 1, \tag{1.20}$$

вследствие чего L'_{jk} удовлетворяют неравенству

$$0 \leqslant L'_{jk} \leqslant 1. \tag{1.21}$$

Упомянем о том, что рекурсия от p'' или величин еще более высокого порядка $p^{(n)}$ к p может зависеть от выбора пути.

Наш подход, который мы только что изложили, не только вводит новое понятие относительной значимости сигнала, но и дает алгоритм для определения p_j, что приводит к некоторым следствиям как теоретического, концептуального, так и практического порядка. При заданной задаче или заданном ансамбле задач, которые призвана выполнить динамическая система, этот алгоритм позволяет нам выбрать тот сигнал, который надлежит отправить, а именно сигнал с наибольшим p_j. Если имеется несколько одинаковых p_j, то безразлично, какой из соответствующих им сигналов будет отправлен. С концептуальной точки зрения мы получаем возможность решать, что делает интересующая нас динамическая система с информацией: уничтожает, сохраняет или порождает ее. Для этого мы используем понятие информации в смысле традиционной теории информации. Но вместо количества информации, оцениваемого на основе относительной частоты символов, мы вводим в ансамбль сигналов относительную значимость, т. е. величины

$$S^{(0)} = -\sum_j p_j \ln p_j, \tag{1.22}$$

$$S^{(1)} = -\sum_k p'_k \ln p'_k, \tag{1.23}$$

где p_j и p'_k были определены выше. Если $\sum_k p'_k = 1$, как всегда предполагается, и $\sum_j p_j < 1$, то имеет место недостаточность информации. Если $\sum_j p_j = 1$, то мы говорим об уничтожении информации при условии, что

$$S^{(1)} < S^{(0)}, \tag{1.24}$$

о сохранении информации при условии, что

$$S^{(1)} = S^{(0)}, \tag{1.25}$$

и о рождении информации при условии, что

$$S^{(1)} > S^{(0)}. \tag{1.26}$$

Смысл такого определения станет сразу же ясен, если мы рассмотрим несколько частных случаев. Например, если к одному и тому же аттрактору приводят два сигнала, то в системе имеется избыточность информации и количество информации (в традиционном техническом смысле слова) становится меньше. Информация убывает от

$$S^{(0)} = -K\left[\frac{1}{2}\ln\left(\frac{1}{2}\right) + \frac{1}{2}\ln\left(\frac{1}{2}\right)\right] = K\ln 2 \tag{1.27}$$

до

$$S^{(1)} = -K \cdot 1 \cdot \ln 1 = 0. \tag{1.28}$$

В случае взаимно однозначного отображения p_j на p'_k мы находим преобразование, которое переводит $\{p_j\}$ в то же самое множество $\{p'_k\}$ с точностью до, быть может, перестановки индексов, т.е. другой нумерации состояний. Ясно, что в этом случае выполняется условие (1.25). Наконец, в случае (1.26), величины p_j, из которых одна равна единице, а все остальные нулю, отображаются, например, на множество величин p, из которых $p' = p'' = {}^1/_2$, а все остальные равны нулю. В этом случае величина $S^{(0)} = -K \cdot 1 \cdot \ln 1 = 0$ возрастает до

$$S^{(1)} = -K\left[\frac{1}{2}\ln\left(\frac{1}{2}\right) + \frac{1}{2}\ln\left(\frac{1}{2}\right)\right] = K\ln 2. \tag{1.29}$$

Разумеется, эти примеры отнюдь не доказывают определений (1.24)–(1.26), а лишь поясняют их смысл.

Наш подход, основанный на синергетике, обладает и рядом других привлекательных черт. Семантика стала проблемой исследования отклика (аттракторов) динамической системы. Система может исправлять ошибки (или служить источником частичной информации). Если приходящий сигнал не переводит начальное состояние **q** *корректно,* т.е. *в аттрактор* (или, если аттрактор не точечный, *на аттрактор*), то он может перевести начальное состояние **q** в *область* притяжения аттрактора, т.е. на склон возвышенности, окружающей дно той самой впадины, которая изображает аттрактор (неподвижную точку). Так система втягивает вектор состояния в аттрактор, соответствующий данной области притяжения, т.е. в *корректное* состояние. Интересная задача состоит в том, чтобы определить минимальное число битов, необходимое для достижения данного аттрактора (или данного уровня относительной значимости).

Предлагаемая нами схема позволяет моделировать и процесс

обучения системы. Система может быть *сенсибилизирована* или *десенсибилизирована* к сигналам *j*; например, может реагировать на конкретный сигнал бо́льшим или меньшим количеством параметров.

В приведенных выше рассуждениях мы исходили из предположения о том, что значимость сигналов измеряется относительно одного и того же приемника. На следующем этапе наших рассуждений допустимо считать, что очередной сигнал застает приемник в *другом* начальном состоянии, чем предыдущий сигнал.

Тем самым мы получаем взаимодействие сигналов и относительную значимость сигналов, зависящую от сигналов, принятых приемником раньше. В общем случае относительная значимость сигнала зависит от последовательности сигналов некоммутативно. Под воздействием сигналов приемник преобразуется снова и снова, и относительная значимость сигналов становится функцией времени.

Еще одно замечание могло бы быть полезно, в особенности в отношении синергетических процессов. Синергетическая система не обязательно должна быть динамической системой, обладающей, например, предельным циклом или хаотическим поведением, но может быть системой, в которой происходят необратимые процессы, приводящие к переходу, например, от неупорядоченного жидкого состояния к структурированному твердотельному состоянию.

Мы хотели бы завершить этот раздел замечанием о распознавании образов, которое будет неоднократно разъясняться с различных точек зрения в этой книге. Распознавание образов можно рассматривать как обработку поступающих сигналов приемником, например мозгом или машиной. В этой связи представляет интерес рассмотреть распознавание образов в духе изложенных выше идей. Я полагаю, что распознавание образов, по крайней мере о общих чертах, представляет собой многоэтапный процесс, в котором активную роль играет приемник. На первом этапе образ воспринимается на глобальном уровне, на котором возможны переходы из начального состояния в несколько аттракторов. Затем включается сенсорная система, позволяющая учесть дополнительные особенности изображения и тем самым выбрать более тонко детализированное множество аттракторов. Более конкретно это означает следующее. На первом этапе воспринимаются, например, общие очертания контурных линий объекта, т. е. устанавливается, что по своим очертаниям он близок к окружности, прямоугольнику и т. д. Затем в случае окружности возникает необходимость дифференцирования нескольких аттракторов: яблоко, человеческое лицо, колесо, дерево. Приемник может запросить о недостающих деталях, например о цвете, наличии вертикальной линии внутри окружности (нос?) и т. д. Так, шаг за шагом, процесс распознавания образов может быть продолжен.

Следует подчеркнуть, что наша интерпретация распознавания образов отличается от традиционного подхода, о котором мы расскажем в гл. 12. При традиционном подходе образ сначала разлагается

на свои *элементарные составляющие,* или *характерные черты.* Мы же начинаем с *глобальной структуры* (контурной линии) и затем переходим ко все более мелким деталям.

Такой подход позволяет нам объяснить (или по крайней мере указать на возможность объяснения), почему при распознавании образов человеком даже прерывистая контурная линия достраивается, и человек «видит» сплошной контур.

1.6.3. Саморождение смысла

Как уже упоминалось, синергетику можно рассматривать как теорию возникновения новых качеств на макроскопическом уровне. При надлежащей интерпретации результатов синергетики мы можем рассматривать *возникновение смысла* как возникновение нового качества системы, или, иначе говоря, как *саморождение смысла.* Чтобы понять, как происходит саморождение смысла, мы намереваемся сравнить физическую систему, а именно лазер, с несколькими модельными системами биологии. Начнем с нескольких общих замечаний о роли информации в биологических системах.

Одна из наиболее поразительных особенностей любой биологической системы – необычайно высокая степень координации между ее отдельными частями. В клетке одновременно и согласованно могут происходить тысячи метаболических процессов. У животных от нескольких миллионов до нескольких миллиардов нейронов и мышечных клеток своим согласованным действием обеспечивает координированные движения, сердцебиение, дыхание и кровообращение. Распознавание образов – процесс в высшей степени кооперативный, равно как и речь и мышление у людей. Совершенно очевидно, что все эти высоко координированные, когерентные процессы становятся возможными только путем обмена информацией, которая должна быть произведена, передана, принята, обработана, преобразована в новые формы информации и должна участвовать в обмене информацией между различными частями системы и вместе с тем, как мы увидим в дальнейшем, между различными иерархическими уровнями. Так мы приходим к непреложному выводу о том, что информация является решающим элементом существования самой жизни.

Понятие информации весьма тонкое, и цель этого раздела нашей книги как раз и состоит в том, чтобы еще более прояснить некоторые его аспекты. Как мы увидим, информация не только связана с пропускной способностью канала связи или с командами, отдаваемыми центральным регулятором отдельным частям системы. Она может также обретать роль своего рода среды, существование которой поддерживается отдельными частями системы – среды, из которой эти части получают конкретную информацию относительно того, как им функционировать когерентно, кооперативно. Именно на этом уровне в игру вступает семантика.

Обратимся сначала к физике. Второе начало термодинамики говорит нам, что в замкнутых системах структуры распадаются и системы становятся все более однородными – по крайней мере на макроскопическом уровне. На микроскопическом уровне может царить полный хаос. Именно по этим причинам информация не может порождаться системами в состоянии теплового равновесия; в замкнутых системах в конце концов устанавливается тепловое равновесие. Но система, находящаяся в состоянии теплового равновесия, не может и *хранить* информацию. Рассмотрим типичный пример – книгу. На первый взгляд может показаться, что она находится в тепловом равновесии – ведь мы даже можем измерить ее температуру. Однако полного теплового равновесия книга достигнет лишь после того, как типографская краска продиффундирует и, расплываясь по каждой странице все больше и больше, распространится по ней,– но тогда текст исчезнет.

Этот простой, но поучительный пример говорит нам о том, что любая память, состоящая из замкнутой системы, никогда не находится в состоянии теплового равновесия, и в каждом конкретном случае неизменно следует задать вопрос о том, *как долго* может храниться информация. Рассмотрим теперь открытые системы, которые поддерживаются вдали от теплового равновесия подводимым к системе потоком энергии и/или материи. Как уже упоминалось, в открытых системах даже в неорганическом мире в результате самоорганизации могут возникнуть специфические пространственные или временные структуры. Примерами могут служить лазер, порождающий когерентное излучение, жидкости, образующие пространственные или временные структуры, или химические реакции, в которых наблюдаются периодические колебания, пространственные спирали или концентрические волны. Даже на этом уровне мы можем говорить о создании или хранении информации. С другой стороны, применительно к таким процессам вряд ли уместны такие слова, как целесообразность, цель и смысл.

Рассмотрим действие лазера несколько подробнее, поскольку это позволит нам ввести терминологию, которая окажется полезной при описании биологических и других систем. В лазере большое число атомов погружены в активную среду, например в такой кристалл, как рубин (рис. 1.28). После накачки извне атомы возбуждаются и могут испускать отдельные цуги световых волн (рис. 1.29). Таким образом, каждый атом испускает сигнал, т. е. создает информацию, переносимую световым полем. В полости лазера испущенные цуги волн могут столкнуться с другим возбужденным атомом, что приведет к усилению испускаемой им волны (рис. 1.30). Следовательно, информация может использоваться для усиления сигнала (рис. 1.31). Так как отдельные атомы могут испускать световые волны независимо друг от друга и так как эти волны могут затем усиливаться другими возбужденными атомами, возникает суперпозиция некоррелирован-

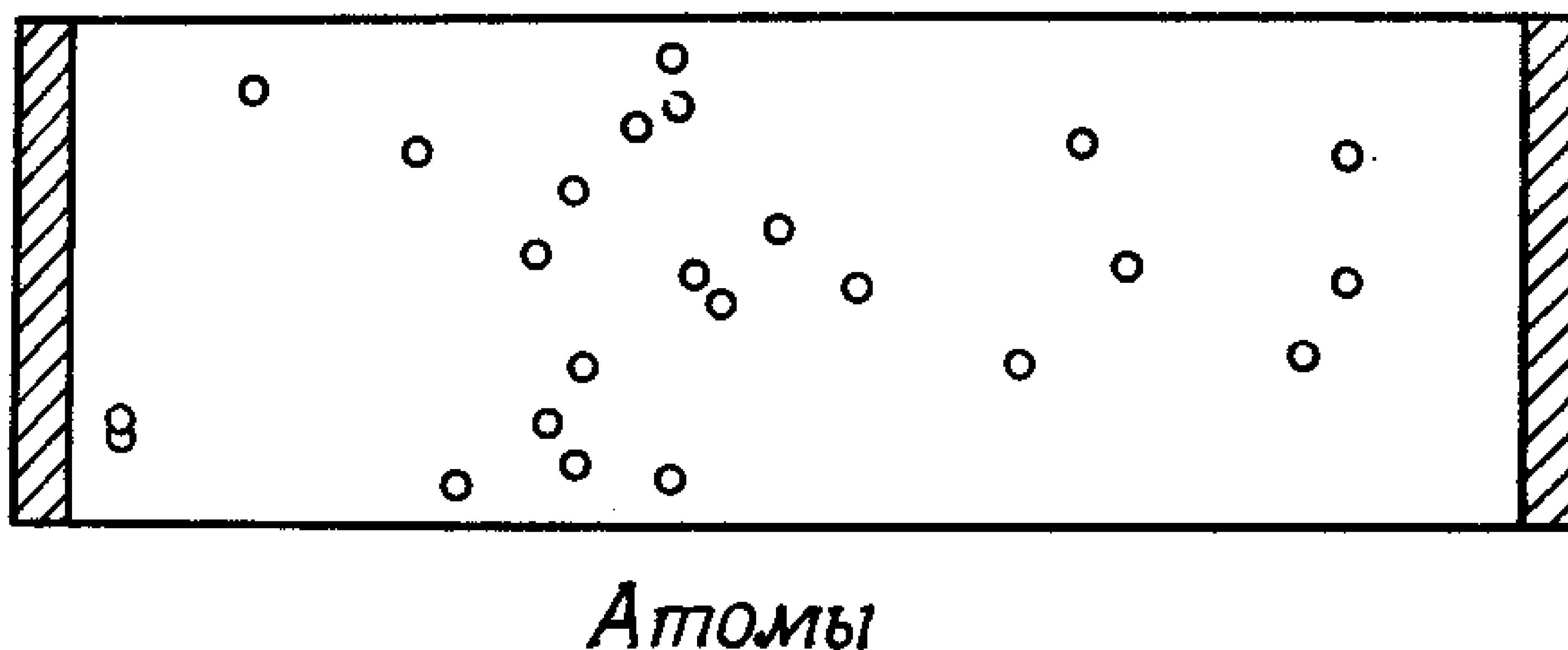


Рис. 1.28. Активные атомы лазера, включенные в кристалл лазерной установки.



Рис. 1.29. Возбужденный атом испускает световую волну (сигнал).

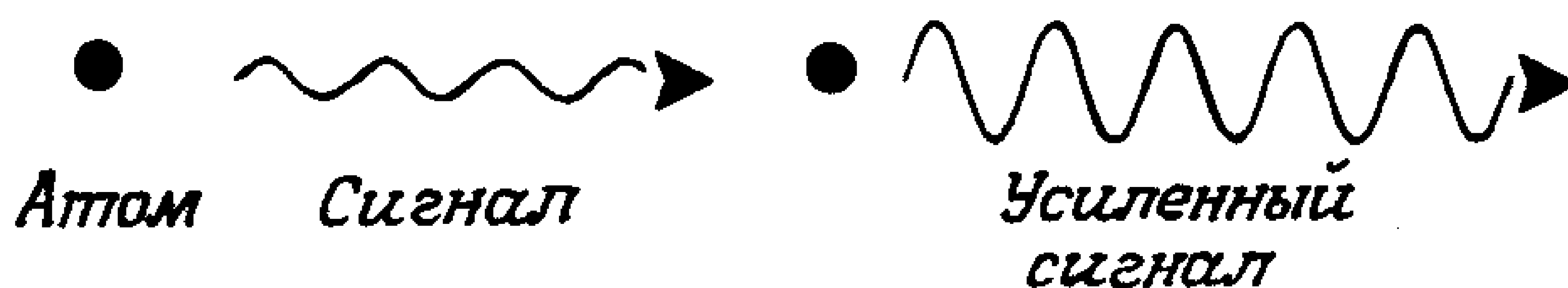


Рис. 1.30. Когда световая волна падает на возбужденный атом, тот может усилить принятый сигнал.

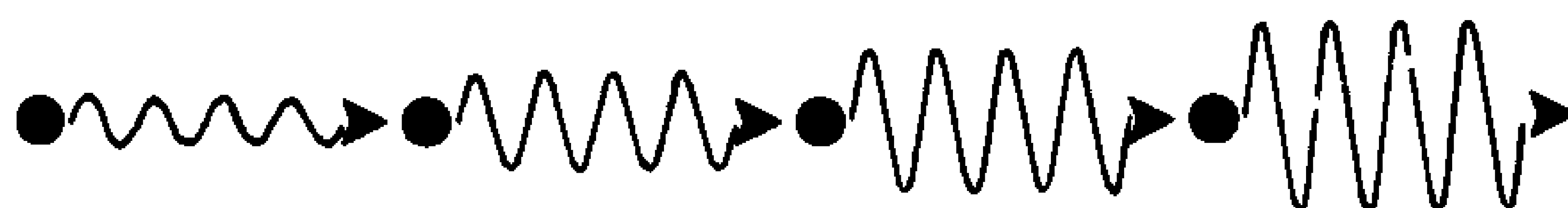


Рис. 1.31. Каскад процессов усиления.

ных, хотя и усиленных цугов волн, и мы наблюдаем совершенно нерегулярную картину (рис. 1.32).

Но, когда амплитуда сигнала становится достаточно большой, начинается совершенно новый процесс. Атомы начинают когерентно осциллировать, и само поле становится когерентным, т. е. оно не состоит более из отдельных некоррелированных цугов волн, а превращается в одну практически бесконечно длинную синусоиду (рис. 1.33).

Перед нами типичный пример самоорганизации: временная структура когерентной волны возникает без вмешательства извне. На смену хаосу приходит порядок. Подробная математическая теория

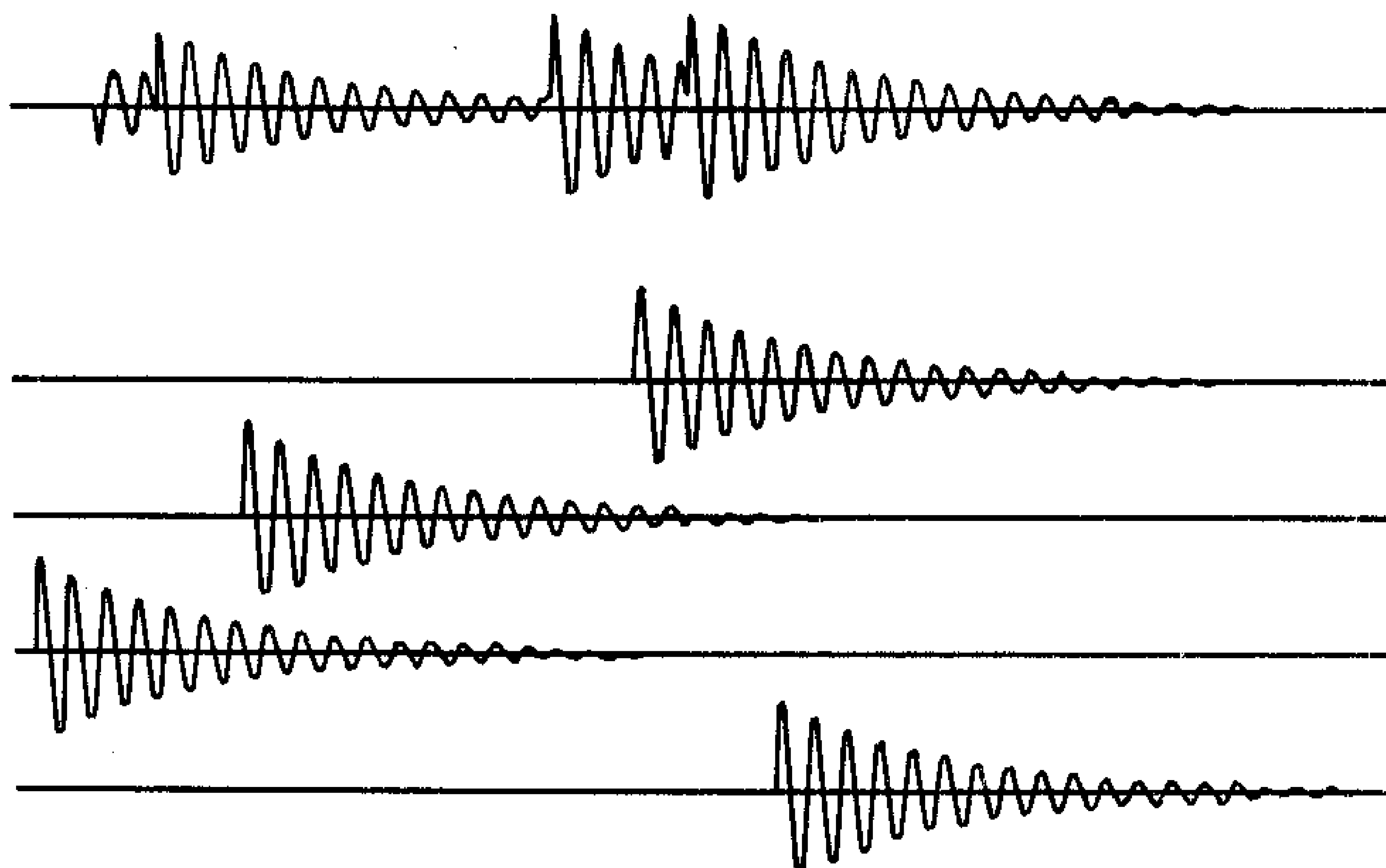


Рис. 1.32. Некогерентная суперпозиция усиленных световых волн приводит к весьма нерегулярному испусканию света.

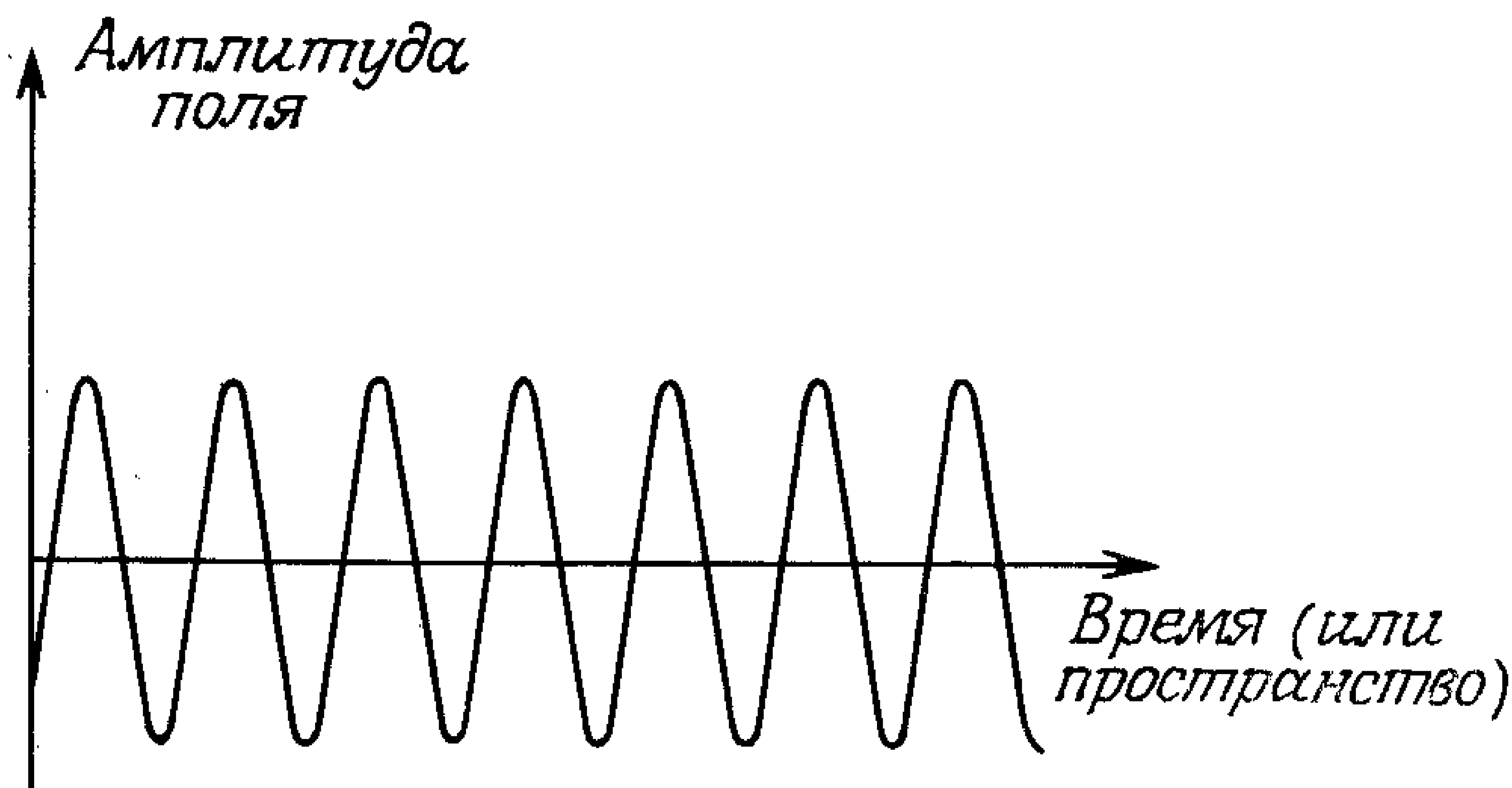


Рис. 1.33. В лазере амплитуда поля имеет вид синусоиды с практически неизменной амплитудой и небольшими флуктуациями фазы.

показывает, что возникающая когерентная световая волна служит своего рода параметром порядка, вынуждающим атомы осциллировать когерентно, или, иначе говоря, подчиняет себе атомы (рис. 1.34). Обращаем внимание читателя на то, что мы имеем здесь дело с циклической причинностью: с одной стороны, параметр порядка подчиняет себе атомы, а, с другой стороны, сам оказывается порожденным совместным действием атомов (рис. 1.35).

С точки зрения информации параметр порядка играет двойную роль: он сообщает атомам о том, как им надлежит вести себя, и, кроме того, доводит до сведения наблюдателя о макроскопически упорядоченном состоянии системы. Если для описания состояний отдельных атомов требуется огромное количество информации, то после установления упорядоченного состояния необходима лишь одна величина, а именно фаза общего светового поля, т. е. происходит

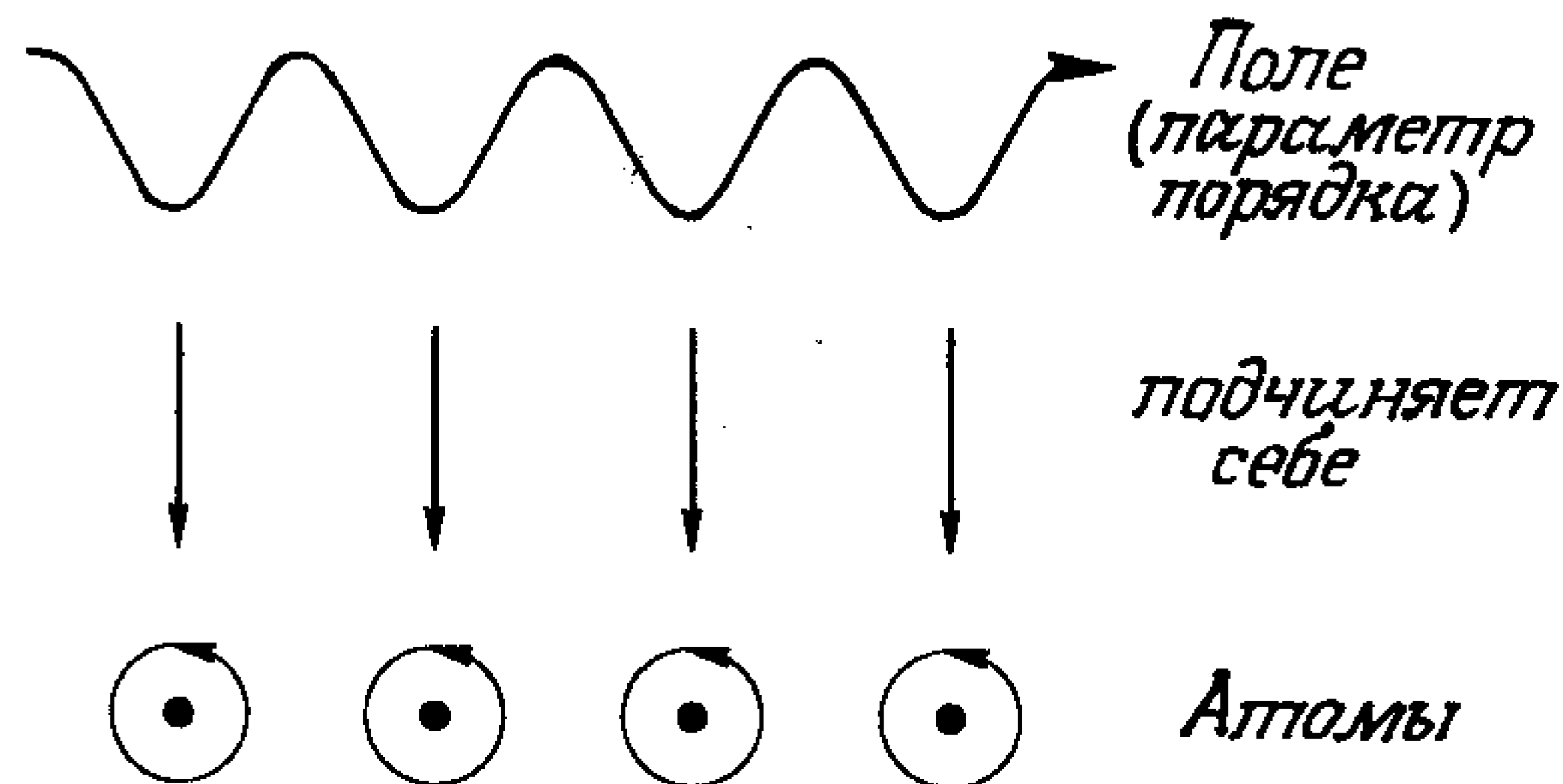


Рис. 1.34. Пример действия принципа подчинения. Поле действует как управляющий параметр и задает характер движения электронов в атомах. Иначе говоря, движение электронов в атомах подчиняется полю.

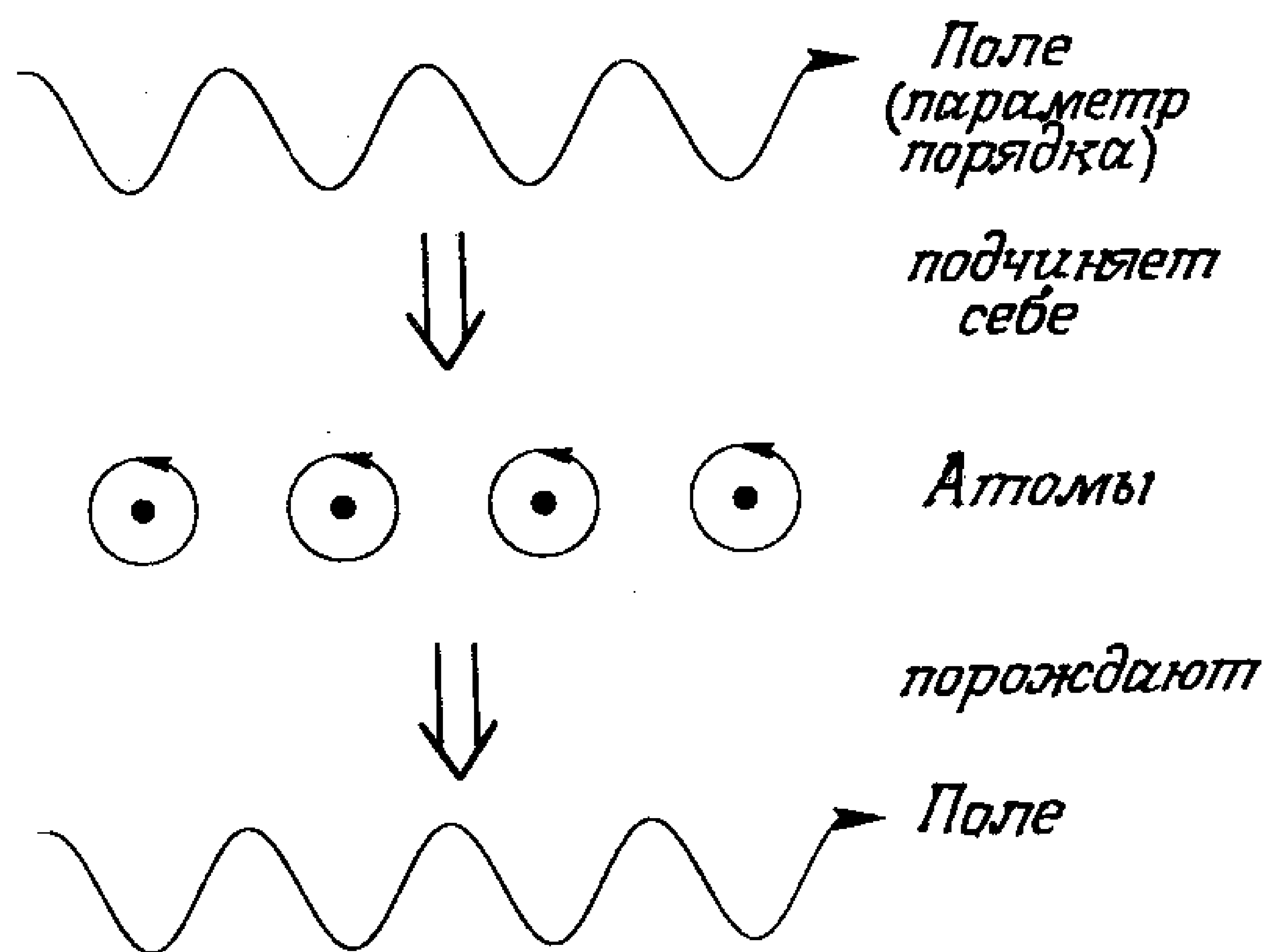


Рис. 1.35. Пример круговой причинности. С одной стороны, поле, действуя как параметр порядка, подчиняет себе атомы. С другой стороны, атомы своим вынужденным излучением порождают поле.

сильное сокращение информации. Мы можем назвать параметр порядка *информатором*. В последние годы было показано, что эти понятия применимы к огромному числу совершенно различных физических, химических и биологических систем.

Чтобы уяснить себе роль обмена информации на ныне рассматриваемом нами уровне, обратимся к примеру – миксомицету Dictiostelium discoideum. Обычно его клетки живут порознь на субстрате, но когда пищевой ресурс исчерпывается, клетки собираются к определенной точке. Механизм такого рода самосборки состоит в следующем.

Отдельные клетки начинают испускать особое вещество – циклический аденозинмонофосфат (цАМФ); таким образом, клетки испускают сигнал, т. е. передают информацию. Когда молекулы цАМФ достигают других клеток, те начинают усиленно вырабатывать

цАМФ так же, как атомы в активной среде лазера усиливают приходящий сигнал (рис. 1.36). Совершенно ясно, что сами клетки «не сознают» смысл получаемой информации, но сложная игра испускания, увеличения концентрации и диффузии молекул цАМФ приводит к образованию спирального распределения концентрации цАМФ, т. е. к рождению информации на более высоком уровне (рис. 1.37). Поскольку эта информация порождается кооперативным действием системы, мы можем назвать ее *синергетической информацией*. Спиральные волны образуют своего рода градиентное поле (информатор), которое отдельные клетки могут изменять и двигаться в сторону наибольшей концентрации в поле. Ясно, что в этом случае мы отчетливо различаем производство информации, носителя информации и приемник информации – клетку, цАМФ и снова клетку. Однако, как нетрудно заметить, на следующем уровне возникает новый смысл, а именно структура распределения концентрации молекул, которая направляет клетки к центру скопления.

По существу та же идея остается в силе, и когда мы переходим к понятию позиционной информации. В этом случае предполагается, что отдельная клетка в ткани получает информацию из химического поля, создаваемого производством и диффузией химического вещества. В общем случае предполагается, что существуют молекулы двух сортов – активаторы и ингибиторы. Там, где существует высокая концентрация молекул-активаторов, специфические гены могут переходить в активное состояние, что и приводит к дифференциации

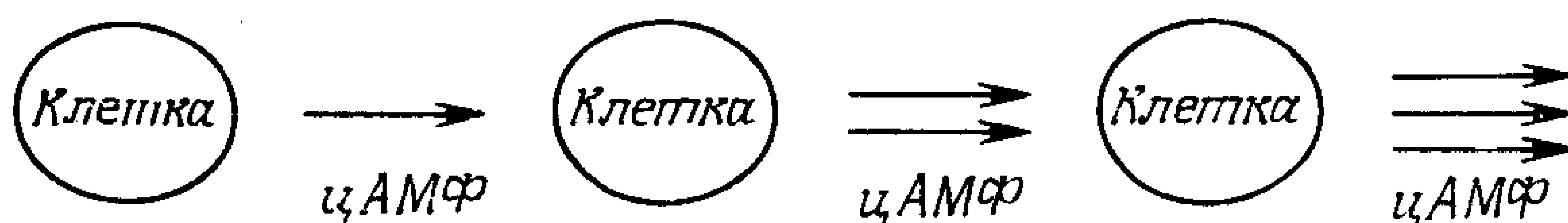


Рис. 1.36. Увеличение числа молекул цАМФ (циклического аденоцилмонофосфата) в клетках миксомицета (усиление).



Рис. 1.37. Схематическое изображение концентрационной волны цАМФ в колонии миксомицетов.

клеток. В этом случае химическое поле играет роль информатора. Конкретной модельной системой была гидра.

Полезно подвести некоторые итоги. Совершенно очевидно, что существует иерархия информационных уровней. На низшем уровне отдельные части системы могут быть источниками информации, которая передается другим частям системы. Передача такой информации может осуществляться между вполне определенными парами элементов или посредством некоторого общего носителя. Примером первого способа передачи информации могут служить аксоны – нервные волокна, соединяющие два нейрона; примерами второго – гормоны, выделяемые в кровь, или феромоны, выделяемые в воздух.

Хотя во всех таких случаях обмен информацией первоначально может происходить случайно, со временем между различного рода сигналами возникает конкуренция или кооперация, и в конце концов устанавливается новое коллективное состояние, которое качественно отличается от неупорядоченного, или некоррелированного, состояния, существовавшего прежде. Это новое состояние описывается одним параметром порядка или набором параметров порядка, или, что эквивалентно, одним или несколькими информаторами. Состояния отдельных частей системы определяются с помощью принципа подчинения. Но тот же процесс можно описать и по-иному – как некоторое специфическое согласование активности отдельных частей системы или как самоорганизацию. Одновременно с переходом в новое состояние происходит сжатие информации. Информация проявляется на макроскопическом уровне и во многих случаях увеличивает надежность и/или эффективность системы или служит другим целям, о которых мы уже упоминали выше.

Новый коллективный уровень становится наблюдаемым для внешнего мира, и при установлении этого отношения, или контекста, достигается новый семантический уровень. Заметим кстати, что отношение, о котором идет речь, может устанавливаться и с внешним миром, что точно так же оно может устанавливаться и внутри самой системы. Это особенно ясно видно на примере с лазером, где кооперативное состояние сопровождается выходом на высокую эффективность. В аналогичном случае с биологической системой коллективное поведение выгодно для всей системы. Находясь вне точек неустойчивости, система может переходить в различные состояния, и ей необходима информация о выборе того или иного состояния. Одна из возможностей состоит в том, что эта информация заложена в системе генетически или поступает в систему от связей, налагаемых одними частями системы на другие. Но нередко в случае вырождения важную роль играет окружающая среда, или, иначе говоря, значимость состояния, в которое может перейти система, оценивается контекстом. По мнению автора, именно так зарождается информация в биологическом смысле. Несмотря на неустойчивость, коллективное состояние возникает, но обретает свой смысл только в соотнесении с

окружающей средой и отчасти в связи с тем, сколь оно важно для выживания всей системы.

Все эти замечания относятся и к генетическому коду, хотя его происхождение пока еще не вполне ясно. Чисто умозрительно можно себе представить, что сначала флуктуации приводят к возникновению некоторых биологических макромолекул, обладающих специфическими свойствами. Наиболее важные из молекул – те, которые обладают способностью размножаться автокаталитически. Ценность информации, передаваемой такой молекулой своему фенотипу, определяется окружающей средой, к которой могут принадлежать другие молекулы и их фенотипы. Возникает взаимодействие между мутацией и отбором новых типов молекул и соответствующих им фенотипов, и тем самым мы становимся свидетелями появления новой информации. Полезна ли эта информация или бесполезна, проверяется только путем взаимодействия новых молекул с окружающей средой.

Приведенные выше соображения относятся к начальным стадиям формирования упорядоченных, или структурированных, коллективных состояний. Но в отличие от упоминавшихся выше физических систем – лазеров, гидродинамики или химических реакций, в биологии появляется новая особенность – отвердевание. Например, когда гены клетки приводятся в активное состояние молекулами-активаторами, клетки дифференцируются, превращаясь в специфические клетки, не поддающиеся дальнейшей модификации и не трансформируемые в исходную клетку. В частности, динамические процессы могут приводить и к образованию твердых структур, таких, как кость или некоторые органы. Аналогичным образом информация отвердевает в ДНК, т. е. в генетическом коде. Насколько можно судить, низшие животные с весьма жесткой «сетью проводов» их нервной системы строятся в основном по правилам, запечатленным в генетическом коде.

С другой стороны, у высших животных помимо жесткой «сети проводов» нервной системы присутствует изрядная доля самоорганизации. Взаимодействие системы с окружающей средой вместе с генетической информацией, запечатленной в системе, приводит к образованию новой информации. В непрестанном апробировании новой информации, хранящейся в мозгу и создаваемой в нем окружающей средой, рождаются новые контексты, и возникает семантика нового рода. Но можно также ожидать, что «отвердевание» происходит на различных иерархических уровнях семантической информации и служит для того, чтобы сделать систему более надежной, и для хранения информации (память). Но если представление о синапсе Хебба, упрочняющемся по мере использования, может быть вполне правильным, то построение семантики требует высокой степени кооперативности внутри системы и повторного взаимодействия с внешним миром. В этом отношении семантическая информация является не статическим свойством, а процессом, в котором кон-

тексты и их соответствие проверяются, усиливаются или отбрасываются снова и снова. Кстати сказать, я убежден, что сознание правильнее считать не статическим состоянием, а процессом, в ходе которого информация непрерывно передается между различными частями мозга и многократно в них обрабатывается.

Здесь, по-видимому, уместно сказать несколько слов о распознавании образов. Низшие животные немедленно реагируют на раздражители, например на вспышки света, и для отклика на сигнал им необходимы лишь весьма немногочисленные критерии, например порог интенсивности. У высших животных поступающая информация заведомо сравнивается с хранящейся информацией. Однако относительно того, как именно происходит это сравнение, наши представления несколько изменяются.

Очень часто предполагается, что поступающий образ сравнивается с некоторыми шаблонами. Но хранение шаблона потребовало бы весьма большого количества информации. Следовательно, можно представить себе картину в духе синергетики и считать, что только наиболее характерные отличительные черты хранятся в виде параметров порядка и могут быть в случае необходимости вызваны для воссоздания целиком всего образа во всех деталях. В этом смысле распознавание образов становится активным процессом, в котором новые образы формируются мозгом в ходе самоорганизации, и мозг, используя определенные гипотезы, многократно проверяет их, сравнивая с поступающими образами. Например, хорошо известно, что, вглядываясь в лица, люди сосредоточивают свое внимание на таких характерных деталях, как глаза, нос или рот, и неоднократно присматриваются к ним.

Наконец, мы не можем не обсудить еще один вопрос, имеющий отношение непосредственно к людям. В отличие от животных люди могут переносить информацию не только с помощью генетического кода, но и с помощью обучения, которое в животном мире происходит лишь в весьма ограниченных пределах. Именно поэтому столь значительная часть нашей культуры опирается на обучение как на новый способ передачи информации от одного поколения к следующему. Но здесь из-за огромного количества знаний, накопленных человечеством, возникают большие трудности. Следовательно, совершенно в духе синергетики важно найти общие унифицирующие идеи и принципы, чтобы справиться со столь огромным количеством информации.

Кроме того, наш подход приводит к картине, существенно отличающейся от тех, которые мы традиционно заимствовали у биологических систем. Принято считать, что существует единый командный центр, например, в мозгу, который организует все поведение. В модели, которую упорно поддерживаем мы, основной акцент делается на процессах самоорганизации, и не так давно нам удалось доказать свою гипотезу с помощью предложенной нами количествен-

ной теории специальных экспериментов по корреляции движений рук и их изменений. В этих экспериментах, выполненных Келсо, испытуемых просили двигать параллельно указательными пальцами. При повышении частоты колебаний наблюдалась самопроизвольная смена параллельных движений на антипараллельные. Ход такого перехода со всеми подробностями прекрасно укладывается в рамки предположения о самоорганизации движения нейронов и мышц.

Это заведомо экстремальный случай, и, вообще говоря, производство и перенос информации в биологических системах надлежит рассматривать двояко: с традиционной точки зрения, согласно которой определенные моторные программы предназначены для выполнения определенных действий, и с новой точки зрения, согласно которой все остальные явления представляют собой не что иное, как различные формы самоорганизации. Имеются основания полагать, что самоорганизация играет в процессах обработки информации, происходящих в биологических системах, главную роль и широко распространена. К такому выводу мы пришли исходя из необычайной гибкости биологических систем, их приспосабливаемости и пластичности.

По моему мнению, исследование информации в биологических системах представляет интерес и в современном обществе, надлежащее функционирование которого основано на адекватном производстве, передаче и переработке информации. Возможно, наиболее важным аспектом, возникшим в последнее время, следует считать круговую причинность, приводящую к коллективному состоянию, которому в социологии могут соответствовать социальный климат, широкое общественное мнение, демократия или диктаторский режим.

1.6.4. Сколько информации необходимо для поддержания упорядоченного состояния?

Обратимся к нашему стандартному примеру, а именно к лазеру. Предположим, что атомы в лазере имеют по два уровня. Общее число атомов в нижнем состоянии обозначим через N_1, число атомов в верхнем состоянии обозначим через N_2. Справедливо соотношение

$$N_1 + N_2 = N. \qquad (1.30)$$

С точки зрения квантовой механики числа заполнения N_1 и N_2 можно связать с вероятностью заполнения

$$p_j = \frac{N_j}{N}, \quad j = 1,\ 2 \qquad (1.31)$$

для отдельного атома. Таким образом, информация, приходящаяся на один атом, составляет величину

$$i = -p_1 \ln p_1 - p_2 \ln p_2, \qquad (1.32)$$

а информация, приходящаяся на все атомы, – величину

$$I = -N(p_1 \ln p_1 + p_2 \ln p_2). \tag{1.33}$$

Как известно, возбужденный атом может испускать фотон либо самопроизвольно (спонтанное испускание), либо, если имеются другие фотоны, вынужденно (так называемое индуцированное испускание). Один фотон можно идентифицировать с символом, переносящим элемент информации. Фотоны могут вылетать из активной среды лазера через зеркала. Следовательно, можно задать вопрос: «Какой должна быть скорость производства фотонов, чтобы поддерживать когерентное состояние?»

Согласно теории лазеров, мы должны в качестве переменных взять не только число фотонов n, но и инверсию, которая определяется как разность между заселенностями верхнего и нижнего состояния:

$$D = N_2 - N_1. \tag{1.34}$$

Согласно теории, скорость производства фотонов определяется уравнением

$$\frac{dn}{dt} = WDn - 2\varkappa n. \tag{1.35}$$

Первый член в правой части описывает скорость производства фотонов (W – константа скорости производства), а второй член учитывает вылет фотонов сквозь зеркала, уменьшающий эффективное производство. Как показано в теории лазеров, формула (1.35) описывает производство когерентных фотонов; производством некогерентных фотонов в теории лазеров пренебрегают. Из-за происходящего в лазере процесса инверсия заселенности также изменяется во времени. Скорость изменения заселенности определяется уравнением

$$\frac{dD}{dt} = \frac{D_0 - D}{T} - 2WDn. \tag{1.36}$$

Здесь D_0 – инверсия, порожденная процессом накачки и процессами релаксации, не приводящими к испусканию света лазером; T – время, за которое любое отклонение инверсии заселенности релаксирует к инверсии D_0. Последний член в уравнении (1.36) связан с лазерным процессом, в котором порождаются фотоны. В общем случае постоянная затухания $\varkappa$ во много раз меньше константы скорости $1/T$. Это позволяет применить так называемое адиабатическое приближение и записать, что

$$\frac{dD}{dt} \approx 0. \tag{1.37}$$

Подставляя соотношение (1.37) в уравнение (1.36), для D получаем

$$D = \frac{D_0}{1 + 2TWn}. \quad (1.38)$$

Когда лазер находится не слишком далеко от порога режима генерации, мы можем разложить знаменатель в степенной ряд по числу фотонов n и, сохраняя только главные члены, получить приближенную формулу

$$D \approx d_0 - 2D_0 TWn. \quad (1.39)$$

Подставляя это выражение в уравнение (1.35) для скорости производства фотонов, преобразуем последнее к виду

$$\frac{dn}{dt} = (WD_0 - 2\varkappa)n - 2TW^2 D_0 n^2. \quad (1.40)$$

Если второй член в скобке (1.40) всегда приводит к уменьшению скорости производства фотонов, то первый обусловливает положительную скорость производства фотонов при условии, что выполняется неравенство

$$WD_0 - 2\varkappa > 0. \quad (1.41)$$

Соотношение (1.41) совпадает с условием генерации лазера. Таким образом, неравенство (1.41) обеспечивает положительную эффективную скорость производства фотонов, или, иначе говоря, *положительную эффективную скорость производства сигналов*. Она необходима для поддержания ненулевого потока фотонов и, следовательно, упорядоченного состояния лазера. Согласно неравенству (1.41), этого можно достичь лишь при достаточно большой инверсии D_0, которая возникает при достаточно сильной накачке. Из выражения

$$n = \frac{WD_0 - 2\varkappa}{2TW^2 D_0} \quad (1.42)$$

мы заключаем, что условие (1.41) обеспечивает ненулевое число фотонов, которые должны при любом t находиться в лазере.

Исследуем теперь, что происходит с информацией (1.32) или (1.33) при увеличении скорости накачки, или, иначе говоря, инверсии D_0. Для этого подставим выражение (1.42) в приближенную формулу (1.39) и получим, что

$$D \approx \frac{2\varkappa}{W} = \text{const}(!) \qquad \text{при } n > 0, \quad (1.43)$$

$$D = D_0 \qquad \text{при } n = 0. \quad (1.44)$$

Иначе говоря, если мы начинаем с низких скоростей накачки, то инверсия D_0 мала и фотонов в лазере нет. Затем D возрастает с такой же скоростью, как D_0. Но когда лазер начинает действовать, инверсия D, как показывает формула (1.43), становится постоянной и

не обнаруживает никаких признаков дальнейшего увеличения. Вся дополнительная энергия, подводимая к лазеру, преобразуется в когерентные фотоны. Из соотношений (1.34) и (1.30) получаем новые соотношения

$$N_1 = \frac{1}{2}(N - D), \tag{1.45}$$

$$N_2 = \frac{1}{2}(N + D), \tag{1.46}$$

которые по формуле (1.31) преобразуются в вероятности заселенности

$$p_1 = \frac{1}{2}\left(1 - \frac{D}{N}\right), \tag{1.47}$$

$$p_2 = \frac{1}{2}\left(1 + \frac{D}{N}\right). \tag{1.48}$$

Подставляя выражения (1.47) и (1.48) в соотношение (1.32), получаем

$$i = -\frac{1}{2}\left(1 - \frac{D}{N}\right)\ln\frac{1}{2}\left(1 - \frac{D}{N}\right) - \frac{1}{2}\left(1 + \frac{D}{N}\right)\ln\frac{1}{2}\left(1 + \frac{D}{N}\right). \tag{1.49}$$

В дальнейшем нам понадобится параметр

$$\gamma = \frac{D}{N}. \tag{1.50}$$

Так как D принимает значения в диапазоне

$$-N, \ldots, +N, \tag{1.51}$$

параметр γ должен быть заключен в интервале

$$-1, \ldots, +1. \tag{1.52}$$

График зависимости i от γ представлен на рис. 1.38. Соотношение (1.49) вместе с соотношениями (1.43) и (1.44) позволяет исследовать изменение информации, приходящейся на отдельный атом, при увеличении скорости накачки D_0. Как следует из рис. 1.39, информация сначала возрастает, достигает максимума и затем выходит на насыщение. В том виде, в каком он существует сейчас, наш подход не в состоянии учесть информацию, содержащуюся в световом поле, поскольку мы не рассматривали никаких флуктуаций, т. е. распределения вероятности по числу фотонов n. Одна из главных целей этой книги состоит в исследовании информации как функции интенсивности накачки не только для атомов, но и для фотонов.

Мы придем к удивительному результату: оказывается, что вблизи точки, где начинается генерация лазерного излучения, информация содержится в фотонах, а не в атомах.

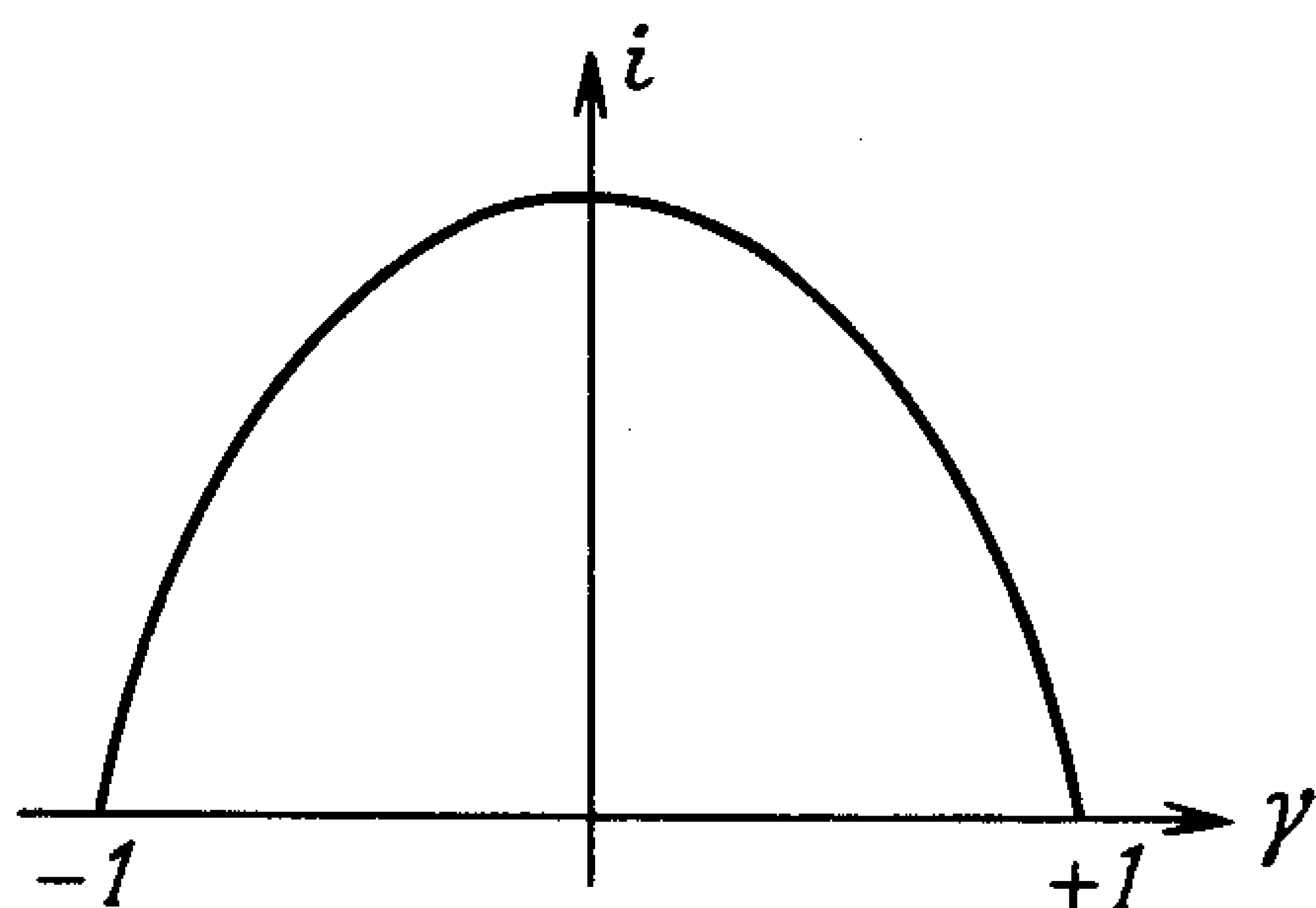


Рис. 1.38. Информация атома как функция параметра γ.

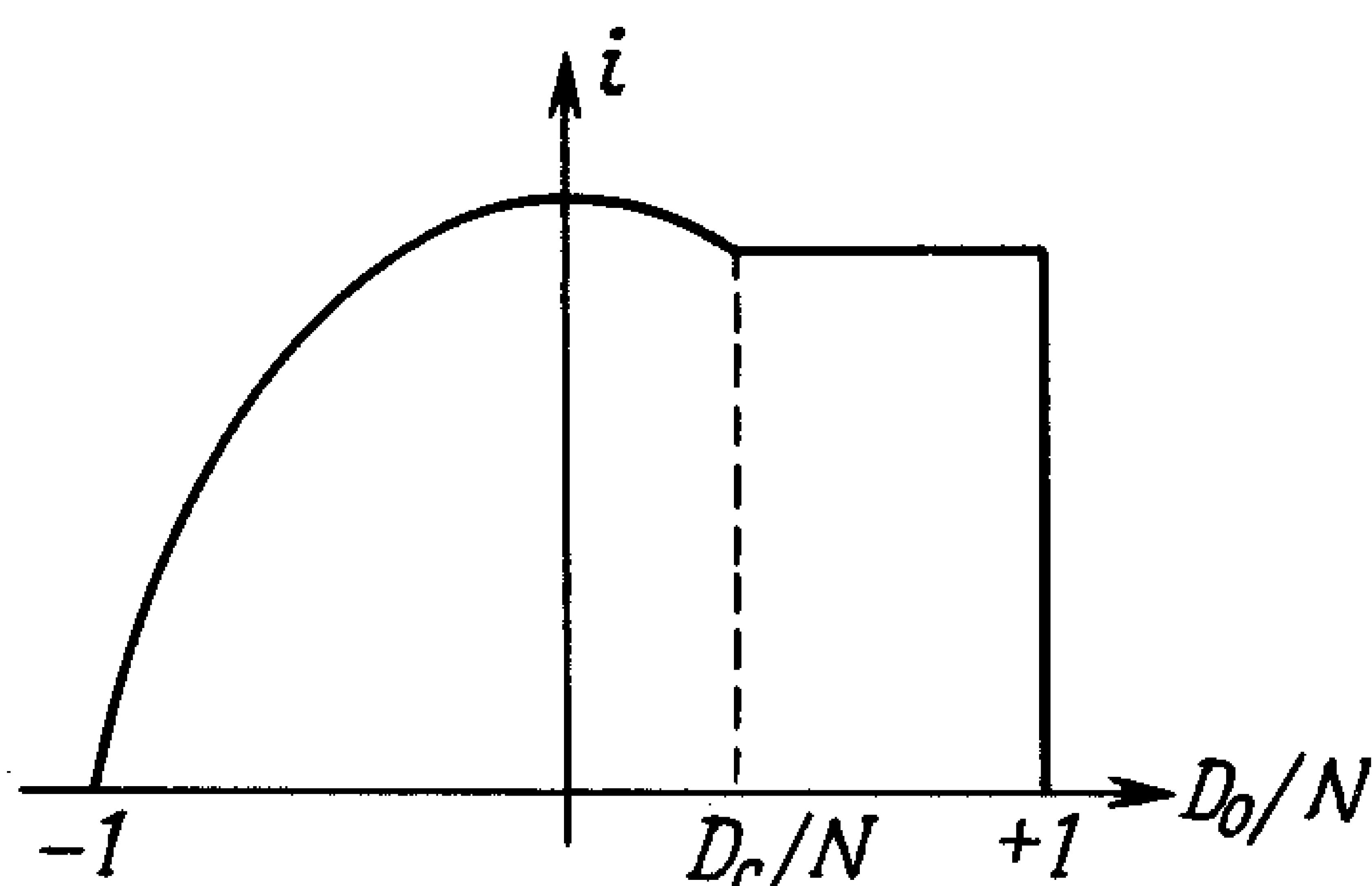


Рис. 1.39. Ход изменения информации отдельного атома в лазере.

1.7. Второе начало синергетики

Обсудив в предыдущих разделах качественные аспекты информации и самоорганизации, мы хотим перейти к самой сути нашего подхода, которому будем неукоснительно следовать в дальнейшем. Начнем с того, что кратко напомним читателю, чем мы до сих пор занимались в области синергетики. Обычно мы выбирали микроскопический или мезоскопический уровень и формулировали свои уравнения. Затем используя такие понятия, как неустойчивость, параметры порядка и подчинение, которым можно придать строгую математическую форму, мы доказывали возможность возникновения структур и тем самым новых качеств на макроскопическом уровне.

В некотором смысле такой подход аналогичен подходу статистической механики. Теперь мы намереваемся развить подход, который можно было бы уподобить подходу термодинамики, а именно мы хотим описывать поведение сложных систем с помощью макроскопически наблюдаемых величин. Затем мы попытаемся угадать микроскопическую структуру процессов, порождающих макроскопическую структуру или макроскопическое поведение. Нашим орудием

для достижения этой цели послужит принцип максимума энтропии, или принцип максимума информационной энтропии, разработанный в совершенно общем виде Джейнсом.

Этот принцип подробно изложен в гл. 3. Пока же достаточно кратко назвать основные идеи. Мы начинаем с макронаблюдаемых, которые могут флуктуировать и имеют известные средние значения. Эти макронаблюдаемые мы различаем по индексу k и обозначаем их средние значения через f_k. Затем мы пытаемся установить распределение вероятностей p_j состояний системы, которым присвоен индекс j. Сделать это можно, если найти максимум информации

$$i = -\sum_j p_j \ln p_j \tag{1.53}$$

при ограничении

$$\sum_j p_j f_j^{(k)} = f_k. \tag{1.54}$$

Ясно, что $f_j^{(k)}$ – это вклад состояния j в макропеременную, которой присвоен индекс k. Кроме того, мы требуем, чтобы

$$\sum_j p_j = 1, \tag{1.55}$$

т.е. чтобы распределение вероятностей p_j было нормировано на единицу. Как было показано Джейнсом и излагается в гл. 4, принцип максимума информационной энтропии позволяет весьма быстро и изящно вывести основные формулы термодинамики. При этом выводе все ограничения относятся к сохраняющимся величинам замкнутой системы, т.е. к энергии, числу частиц и т.д. Трудность проблемы обобщения этого принципа максимума энтропии на системы, далекие от теплового равновесия, или даже на нефизические системы, кроется именно в адекватном выборе ограничений.

Как мы увидим, использовавшиеся до сих пор ограничения, например сохранение энергии или даже регулируемых потоков энергии, поступающих в систему, неадекватны для рассмотрения открытых систем и особенно перехода от бесструктурного состояния к структурированному, как это происходит при неравновесных фазовых переходах. Принцип максимума энтропии неоднократно подвергается критике, поскольку выбор ограничений, по утверждению оппонентов, вносит определенную субъективность и ограничения выбираются произвольно по воле наблюдателя, а не в соответствии с объективными критериями.

Джейнс подробно проанализировал такого рода критику и выдвинул контраргументы. Я же со своей стороны хочу высказать еще одну точку зрения. Научный прогресс опирается на общий консенсус, достигаемый в научном сообществе; результаты признаются объек-

тивными с общего согласия. Такой подход можно было бы назвать «относительным объективизмом». В действительности относительный объективизм – максимум того, что можно сказать о любой физической теории, поскольку в естественных науках теорию нельзя доказать – ее можно только опровергнуть. Это совершенно правильно отметил Поппер. Таким образом, нам не остается ничего другого, как принять процесс обучения на основе правильного выбора адекватных ограничений. Именно так и происходит в термодинамике, где, насколько известно, адекватными ограничениями являются законы сохранения.

В области неравновесных фазовых переходов или, если говорить более общо, открытых систем мы хотим сделать первые шаги и продемонстрировать, в чем состоят такие ограничения. Разумеется, если мы ограничиваем свой анализ неравновесными фазовыми переходами, то во всех случаях, когда известны микроскопические функции распределения, обнаруживается полное согласие между макроскопическим подходом на основе принципа максимума (информационной) энтропии и результатами микроскопической теории. Поэтому я уверен, что и в нашей области можно достичь консенсуса.

Существует другой аспект, важный с математической точки зрения. Дело в том, что, когда мы накладываем на систему определенные ограничения, заданные экспериментальными средними значениями, принцип максимума (информационной) энтропии всегда приводит к функции распределения, воспроизводящей эти средние значения. В этом смысле мы имеем здесь дело с тавтологией. Однако на следующем шаге мы можем ввести новые средние значения с помощью распределения вероятностей и затем делать предсказания, допускающие экспериментальную проверку. Если наши предсказания не сбываются, то мы можем выбирать в качестве дополнительных ограничений новые экспериментальные данные, которые порождают новые функции распределения вероятностей. Этим мы начинаем бесконечный процесс. Но, несмотря на столь предостерегающее замечание, мы можем достичь консенсуса относительно правильного выбора ограниченного набора ограничений, если мы ограничим свой анализ определенными классами явлений.

Один из таких классов образуют, как уже упоминалось, замкнутые (термодинамические) системы при надлежащих ограничениях. Другой класс состоит из неравновесных фазовых переходов, о которых пойдет речь в нашей книге. Как мы увидим из дальнейшего, этот класс включает в себя многочисленные явления в различных областях, например возникновение пространственных структур, новых типов информации и даже колебательных явлений. Адекватный выбор ограничений для процессов, приводящих к детерминистическому хаосу, остается, по крайней мере частично, задачей будущего.

Как мы увидим, суть того нового, что нам удалось понять с помощью нашего подхода, сводится к следующему. В качестве

первого шага можно высказать предположение о том, что адекватные ограничения должны включать макроскопические переменные, или, иначе говоря, параметры порядка. Но известно, что при неравновесных фазовых переходах происходят критические флуктуации параметров порядка, т.е. флуктуации параметров порядка становятся макроскопическими переменными. Действительно, оказывается, что включение флуктуаций параметра порядка – решающий шаг в нахождении адекватных ограничений для этого класса явлений.

Используя результаты классической теории в качестве руководства, мы на следующем шаге можем сделать нечто гораздо большее, а именно с самого начала действовать без параметров порядка. Вместо параметров порядка мы выбираем в качестве отправного пункта корреляционные функции, т.е. моменты наблюдаемых переменных, по которым мы можем затем реконструировать параметры порядка и подчиненные моды. Кстати сказать, мы можем также восстанавливать макроскопические структуры, или, иначе говоря, автоматизировать распознавание нестационарных образов, возникающих при неравновесном фазовом переходе.

Завершая наш обзор, обратимся к обсуждению соотношения между аналитическим (или микроскопическим) подходом и холистическим (или макроскопическим) подходом и позволим себе сделать еще одно замечание в пользу макроскопического подхода. Очень часто даже подсистемы становятся настолько сложными, что выписать в явном виде микроскопические или мезоскопические уравнения трудно или даже невозможно. Если взять экстремальный случай – головной мозг человека, то подсистемы, например нервные клетки (нейроны), сами являются необычайно сложными системами. Нервная клетка содержит сому, аксон и до 80 тысяч дендритов, которыми она связана с другими нервными клетками. У человеческого мозга существует около 10 млрд. нервных клеток. Синергетический подход предполагает, что, несмотря на столь чудовищную сложность, ряд схем («паттернов») поведения может быть описан с помощью понятия параметра порядка. Уравнения для параметра порядка в настоящее время устанавливаются на феноменологической основе.

Недавно нам удалось найти удачный пример, а именно координацию движений рук и особенно непроизвольные изменения движений рук. Хотя соответствующие подсистемы весьма многочисленны и состоят из нервных клеток, мышечных клеток и других тканей, поведение (движение рук) может быть описано одним параметром порядка. Ниже мы приводим описание этих экспериментов и поясняем на этом примере некоторые особенности нашего общего подхода. Другие примеры заимствованы нами из лазерной физики и гидродинамики. Анализируя многочисленные примеры, рассматриваемые в синергетике с микроскопической точки зрения, нетрудно указать еще больше приложений нашего нового макроскопического подхода.

Кроме того, нетрудно указать многочисленные примеры, особенно в биологии, в которых применим только макроскопический подход.

В заключение мы хотим заметить, что выражение «принцип максимума информации» мы будем использовать как равнозначное выражению «принцип максимума энтропии». Но, как станет ясно из дальнейшего, термин «информация» лучше подходит к ситуации в неравновесных системах.

Глава 2
От микроскопического мира к макроскопическому

2.1. Уровни описания

В этой главе я изложу основные понятия и методы, которые использовались синергетикой при исследовании самоорганизации с помощью микроскопического подхода. Читатели, знакомые с этим подходом, могут пропустить эту главу и переходить непосредственно к гл. 3. Тем же читателям, кто не знаком с этими понятиями и методами и кто хотел бы глубже овладеть ими, я советую обратиться к моим книгам «Синергетика» и «Синергетика. Иерархии неустойчивостей», в которых эти понятия объяснены весьма подробно.

Когда нам приходится иметь дело с какой-нибудь системой, мы прежде всего идентифицируем те переменные, или величины, посредством которых нам хотелось бы описать систему. Такое описание может быть выполнено на различных уровнях, которые, однако, взаимосвязаны. Рассмотрим в качестве примера жидкость (рис. 2.1). На микроскопическом уровне жидкость может быть описана как состоящая из отдельных молекул. Следовательно, для полного описания жидкости нам понадобятся радиусы-векторы и скорости отдельных молекул. Однако для многих целей достаточно рассматривать

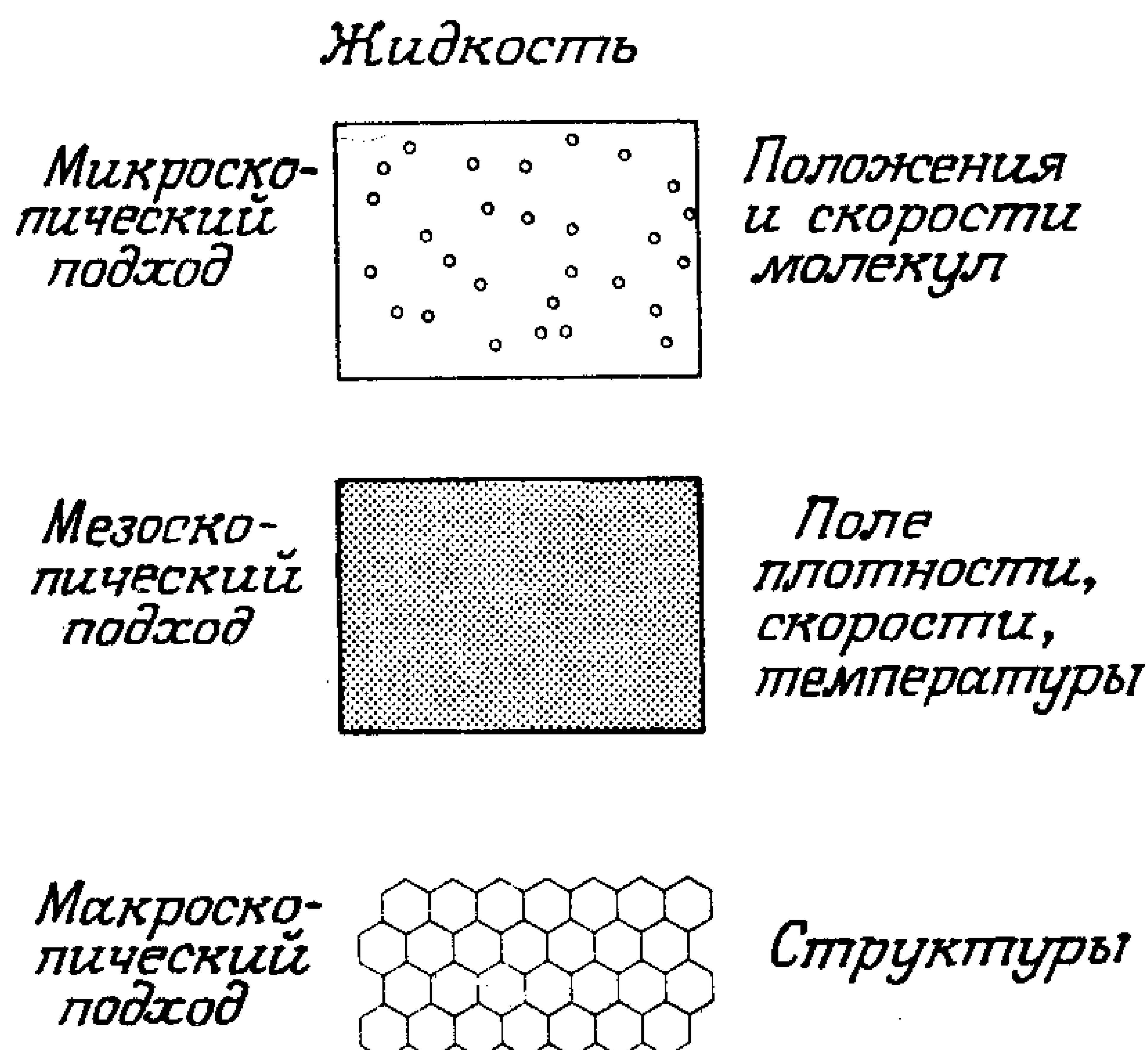


Рис. 2.1. Схема микроскопического, мезоскопического и макроскопического подходов на примере жидкости.

Таблица 2.1. Примеры мезоскопических переменных

Область исследования	Переменные
Химические реакции, отвердевание	Плотности молекул в различных фазах
Жидкости	Поля скоростей
Пламена	Поля температур
Плазма	Электрическое и магнитное поля
Лазеры, параметрические генераторы	Поляризация атомов, инверсия
Физика твердого тела, диод Ганна, нитевидные кристаллы	Плотности электронов и дырок
Морфогенез	Плотности клеток в тканях
Предбиотическая эволюция	Число биомолекул
Популяционная динамика	Число особей
Нейронные сети	Скорости возбуждения нейронов
Движение	Растяжение и сокращение мышц
Экономика	Денежные потоки
Социология	Число людей, разделяющих определенные взгляды
Синергетические компьютеры	Активация элементов

жидкость на мезоскопическом уровне. В этом случае мы начинаем с элементов объема, которые еще малы по сравнению с полным объемом жидкости, но все же достаточно велики для того, чтобы мы могли спокойно говорить о плотностях, полях скоростей или локальных температур. Наконец, нас может интересовать макроскопический уровень, на котором мы хотим исследовать образование структур, например образование шестиугольных ячеек в жидкости, подогреваемой снизу. Аналогичное разбиение уровней может быть проведено и в случае биологических систем. Но, как легко догадается читатель, в биологии произвол в выборе уровней еще больше.

Рассмотрим пример биологической ткани. На микроскопическом уровне мы можем говорить о биомолекулах, органеллах и т. д. На мезоскопическом уровне мы можем говорить о клетках и, наконец, на макроскопическом уровне мы можем рассматривать целые ткани или органы, состоящие из этих клеток. Совершенно очевидно, что при переходе от мезоскопического уровня к макроскопическому мы должны пренебречь многими мелкими подробностями строения клеток и их составных частей. Вместо мелких деталей мы сосредоточиваем внимание на тех особенностях строении клеток, которые существенны для образования органов. Следовательно, уже на мезоскопическом уровне мы имеем дело с колоссальным сокращением информации.

В этой главе нас будет интересовать главным образом переход от мезоскопического уровня к макроскопическому, хотя в ряде случаев

возможен и прямой переход от микроскопического уровня к макроскопическому уровню. В физике мезоскопический уровень может быть достигнут с помощью статистической механики, где вводятся и рассматриваются некоторые соответствующие существу дела переменные. Однако в большинстве случаев, например в химии и биологии, мы используем феноменологические уравнения для соответствующих переменных.

Чтобы читатель мог составить представление о разнообразии возникающих здесь задач, мы приведем ряд примеров мезоскопических переменных (табл. 2.1).

2.2. Уравнения Ланжевена

Совершенно типичным примером описания на мезоскопическом уровне может служить броуновское движение частицы, например пылинки, погруженной в жидкость. Движение такой частицы описывается уравнением Ланжевена, в котором переменная q идентифицируется со скоростью частицы. Микроскопическое движение всех молекул жидкости порождает два эффекта. С одной стороны, оно приводит к затуханию скорости, а с другой стороны,–к случайным импульсам, испытываемым частицей со стороны жидкости. Уравнение Ланжевена имеет вид

$$\dot{q} = K(q) + F(t), \tag{2.1}$$

где в случае броуновской частицы

$$K(q) = -\gamma q. \tag{2.2}$$

В синергетике часто приходится рассматривать величину $K(q)$, задаваемую нелинейным выражением

$$K(q) = \alpha q - \beta q^3. \tag{2.3}$$

Флуктуирующие силы характеризуются следующими свойствами:

$$\langle F(t) \rangle = 0, \tag{2.4}$$

$$\langle F(t) F(t') \rangle = Q\delta(t - t'), \tag{2.5}$$

где среднее берется по стохастическому процессу. Если мы имеем несколько переменных, объединенных в вектор состояния $\mathbf{q}$, то уравнение Ланжевена принимает вид

$$\dot{\mathbf{q}} = \mathbf{K}(\mathbf{q}) - \mathbf{F}(t) \tag{2.6}$$

и флуктуирующие силы по предположению обладают следующими свойствами:

$$\langle F_j(t) \rangle = 0, \tag{2.7}$$

$$\langle F_j(t) F_{j'}(t') \rangle = Q_j \delta_{jj'} \delta(t - t'). \tag{2.8}$$

Заметим, что $\mathbf{q}$ может быть вектором в многомерном пространстве и тем самым представлять очень сложную систему. Иногда флуктуирующие силы сами зависят от переменной состояния $\mathbf{q}$. Тогда возникает ряд специфических задач, которые могут быть решены с помощью исчислений Ито и Стратоновича. В рамках *подхода Ито* уравнение Ланжевена (2.1) следует заменить уравнением

$$dq(t) = K(q(t))\,dt + g(q(t))\,dw(t). \tag{2.9}$$

Здесь K и g, вообще говоря, нелинейные функции от q, а dw – стохастический процесс, относительно которого нами приняты следующие предположения:

$$\langle dw \rangle = 0, \tag{2.10}$$

$$\langle dw(t)\,dw(t') \rangle = \delta(t - t')\,dt. \tag{2.11}$$

Формулы Ито исходят из допущения о том, что стохастические процессы $q(t)$ и dw, входящие в последний член уравнения (2.9), статистически некоррелированы. Когда мы описываем систему многомерным вектором состояния с компонентами $q_1, \ldots, q_m$, уравнение Ито имеет вид

$$dq_l(t) = K_l(\mathbf{q}(t))\,dt + \sum_m g_{lm}(\mathbf{q}(t))\,dw_m(t), \tag{2.12}$$

где величина dw_m, описывающая стохастический процесс, обладает следующими свойствами:

$$\langle dw_m \rangle = 0, \tag{2.13}$$

$$\langle dw_m(t)\,dw_l(t') \rangle = \delta_{lm}\delta(t - t')\,dt. \tag{2.14}$$

Чтобы продемонстрировать отличительные особенности процедуры Ито, рассмотрим произвольную дифференцируемую функцию

$$\mathbf{u} = \mathbf{u}(\mathbf{q}) \tag{2.15}$$

и ее дифференциал. Из-за свойства (2.14) нам необходимо дойти до второй производной с помощью разложения

$$du_j = \sum_k \frac{\partial u_j}{\partial q_k}\,dq_k + \frac{1}{2}\sum_{kl} \frac{\partial^2 u_j}{\partial q_k \partial q_l}\,dq_k dq_l. \tag{2.16}$$

Подставляя правую часть уравнения (2.12) в разложение (2.16) и сохраняя все члены, содержащие dt, но пренебрегая всеми членами более высокого порядка, приходим к соотношению

$$du_j = \sum_k \frac{\partial u_j}{\partial q_k}\left[K_k(q)\,dt + \sum_m g_{km}\,dw_m(t) + \frac{1}{2}\sum_{kl} \frac{\partial^2 u_j}{\partial q_k \partial q_l}\left[\sum_{mn} g_{km} g_{ln}\,dw_m dw_n\right].\right. \tag{2.17}$$

Напомним теперь читателю кратко, в чем состоит *подход Стратоновича*. Стохастическое уравнение для одной переменной имеет вид

$$dq = K(q)\,dt + g(q)\,dw(t), \tag{2.18}$$

а для компонент вектора состояния

$$dq_l = K_l(\mathbf{q})\,dt + \sum_m g_{lm}(\mathbf{q})\,dw_m(t). \tag{2.19}$$

В отличие от исчисления Ито последние члены в уравнениях (2.18) или (2.19) у Стратоновича интерпретируются иначе, а именно их надлежит вычислять в центре интервала, т. е. заменять $g(q(t_i))\,dw(t_i)$ величиной

$$g\left(q\left(\frac{t_i + t_{i-1}}{2}\right)\right) dw(t_i). \tag{2.20}$$

Иначе говоря, в исчислении Стратоновича q и dw перестают быть статистически независимыми. Как будет показано дальше, мы можем получить уравнение Ито с помощью макроскопического подхода.

2.3. Уравнение Фоккера–Планка

Во многих приложениях, особенно в тех случаях, когда задачи нелинейны, т. е. когда K нелинейно зависит от q, удобно перейти к уравнению Фоккера–Планка, записанного для функции распределения $f(q, t)$. Оно описывает вероятность найти переменную q в интервале $q \to q + dq$ в момент времени t.

Уравнение Фоккера–Планка, соответствующее уравнению Ланжевена (2.1), имеет вид

$$\frac{\partial t}{\partial t} = -\frac{\partial}{\partial q}(K(q)f) + \frac{Q}{2}\frac{\partial^2}{\partial q^2}f, \tag{2.21}$$

где первый член в правой части называется дрейфовым, а второй диффузионным.

Нетрудно найти стационарное решение, определяемое условием

$$\frac{\partial f}{\partial t} = 0, \tag{2.22}$$

если выполняется граничное условие

$$f(q) \to 0 \qquad \text{при} \qquad q \to \pm\infty. \tag{2.23}$$

Это стационарное решение определяется выражением:

$$f = N \exp\left[-\int_{q_0}^{q} \frac{2K(q')}{Q}\,dq'\right], \tag{2.24}$$

где N – нормировочный множитель, такой, что интеграл от f равен единице. В случае многомерного вектора состояния уравнение Фоккера–Планка необходимо записать для

$$f(\mathbf{q}, t). \tag{2.25}$$

В этом случае оно преобразуется к виду

$$\frac{\partial f}{\partial t} = -\sum_j \frac{\partial}{\partial q_j}(K_j f) + \frac{1}{2}\sum_{jk} Q_{jk} \frac{\partial^2}{\partial q_j \partial q_k} f. \tag{2.26}$$

Получить решения уравнения (2.26) в явном виде удается лишь в частных случаях, например, если величина K линейна по переменным q, а величина Q_{jk} не зависит от q. При таких K и Q и зависящие от времени, и стационарные решения удается построить в аналитически замкнутом виде (см. мою книгу «Синергетика. Иерархии неустойчивостей»). Стационарное решение уравнения (2.26) может быть получено в явном виде и в том случае, если выполняется так называемое условие детального равновесия (см. разд. 2.4).

Принимая за исходное *дифференциальное уравнение Ито,* мы приходим к уравнению Фоккера–Планка:

$$\frac{\partial f}{\partial t} = -\sum_k \frac{\partial}{\partial q_k}[K_k(\mathbf{q})f] + \frac{1}{2}\sum_{kl} \frac{\partial^2}{\partial q_k \partial q_l}\left[\sum_m q_{km} q_{lm} f\right]. \tag{2.27}$$

Что же касается *дифференциального уравнения Стратоновича,* то ему соответствует уравнение Фоккера–Планка

$$\frac{\partial f}{\partial t} = -\sum_l \frac{\partial}{\partial q_l}\left\{\left[K_l(\mathbf{q}) + \frac{1}{2}\sum_{kj} \frac{\partial g_{lj}}{\partial q_k} g_{kj}\right]\right\} +$$
$$+ \frac{1}{2}\sum_{lm} \frac{\partial^2}{\partial q_l \partial q_m}\left(\sum_i g_{li} g_{mi} f\right). \tag{2.28}$$

2.4. Точное стационарное решение уравнения Фоккера–Планка для систем, находящихся в детальном равновесии

В этом разделе мы покажем, что при условии детального равновесия стационарное решение уравнения Фоккера–Планка может быть найдено в явном виде с помощью квадратур.

В то время как принцип детального равновесия должен выполняться практически для всех систем, находящихся в состоянии теплового равновесия, он вполне может нарушаться для систем, далеких от теплового равновесия. Таким образом, каждый отдельный случай требует особого рассмотрения (например, привлечения соображений симметрии), и сказать заранее, будет ли применим принцип детального равновесия, просто невозможно. Именно исследование струк-

туры уравнения Фоккера–Планка позволяет решить, выполняется ли принцип детального равновесия.

2.4.1. *Детальное равновесие*

Обозначим через $\mathbf{q}$ набор переменных $q_1, \ldots, q_N$, а набор переменнных при обращении времени обозначим через

$$\tilde{\mathbf{q}} = \{\varepsilon_1 q_1, \ldots, \varepsilon_N q_N\}, \tag{2.29}$$

где $\varepsilon_i = -1$ (или $+1$) в зависимости от того, изменяет (или не изменяет) знак координата q_i при обращении времени. Кроме того, пусть λ – набор параметров, определяемых извне. Набор тех же параметров после обращения времени обозначим через

$$\tilde{\lambda} = \{\nu_1 \lambda_1, \ldots, \nu_M \lambda_M\}, \tag{2.30}$$

где $\nu_i = -1$ (или $+1$) в зависимости от инверсионной симметрии внешних параметров при обращении времени. Обозначим совместную вероятность найти систему в момент времени t_1 в точке с координатами $\mathbf{q}$, а в момент времени t_2 – в точке с координатами $\mathbf{q}'$ через

$$P(\mathbf{q}', \mathbf{q}; t_2, t_1). \tag{2.31}$$

В дальнейшем мы будем рассматривать стационарную систему, поэтому совместная вероятность зависит только от разности времен $t_2 - t_1 = \tau$. Следовательно, вероятность (2.31) можно записать в виде

$$\mathbf{P}(\mathbf{q}', \mathbf{q}; t_2, t_1) = W(\mathbf{q}', \mathbf{q}; \tau). \tag{2.32}$$

Сформулируем теперь принцип детального равновесия. Возможны следующие два определения.

1) Принцип детального равновесия (первый вариант):

$$W(\mathbf{q}', \mathbf{q}; \tau, \lambda) = W(\tilde{\mathbf{q}}, \tilde{\mathbf{q}}'; \tau, \tilde{\lambda}). \tag{2.33}$$

Совместную вероятность можно выразить через стационарное распределение $f(\mathbf{q})$, умноженное на условную вероятность $\mathbf{P}$. Стационарность подчеркивается при этом следующей формой записи:

$$\mathbf{P} = \mathbf{P}(\mathbf{q}' \mid \mathbf{q}; \tau, \lambda). \tag{2.34}$$

Это позволяет записать принцип детального равновесия в несколько ином виде:

$$\mathbf{P}(\mathbf{q}' \mid \mathbf{q}; \tau, \lambda) f(\mathbf{q}, \lambda) = \mathbf{P}(\tilde{\mathbf{q}} \mid \tilde{\mathbf{q}}'; \tau, \tilde{\lambda}) f(\tilde{\mathbf{q}}', \tilde{\lambda}). \tag{2.35}$$

Здесь и далее мы предполагаем, что уравнение Фоккера–Планка имеет единственное стационарное решение. Можно непосредственно показать, что

$$f(\mathbf{q}, \lambda) = f(\tilde{\mathbf{q}}, \lambda). \tag{2.36}$$

Определим вероятность перехода в секунду как величину

$$w(\mathbf{q}',\mathbf{q};\lambda) = [(d/d\tau)\mathbf{P}(\mathbf{q}'\,|\,\mathbf{q};\tau,\lambda)]_{\tau=0}. \qquad (2.37)$$

Взяв производную по τ от обеих частей соотношения (2.35) и полагая $\tau = 0$ (но $\mathbf{q} \neq \mathbf{q}'$), мы придем к соотношению (2.38).

2) Принцип детального равновесия (второй вариант):

$$w(\mathbf{q}',\mathbf{q};\lambda)f(\mathbf{q},\lambda) = w(\tilde{\mathbf{q}},\tilde{\mathbf{q}}';\tilde{\lambda})f(\tilde{\mathbf{q}}',\tilde{\lambda}). \qquad (2.38)$$

У соотношения простой смысл. Его левая часть описывает полную скорость перехода из состояния $\mathbf{q}$ в новое состояние $\mathbf{q}'$. Следовательно, принцип детального равновесия требует, чтобы эта скорость перехода была равна скорости обратного перехода с $\mathbf{q}'$ и $\mathbf{q}$ для обратного движения, например для движения с обращенными импульсами. Можно показать, что первый и второй варианты принципа детального равновесия эквивалентны.

2.4.2. Требуемая структура уравнения Фоккера–Планка и его стационарное решение

Используя условную вероятность $\mathbf{P}$ (которая есть не что иное, как функция Грина), запишем уравнение Фоккера–Планка:

$$\frac{d}{d\tau}\mathbf{P}(\mathbf{q}'\,|\,\mathbf{q};\tau,\lambda) = L(\mathbf{q}',\lambda)\,\mathbf{P}(\mathbf{q}'\,|\,\mathbf{q};\tau,\lambda), \qquad (2.39)$$

где предполагается, что оператор L имеет вид

$$L(\mathbf{q}) = -\sum_i \frac{\partial}{\partial q_i} K_i(\mathbf{q},\lambda) + \frac{1}{2}\sum_{ik} \frac{\partial^2}{\partial q_i \partial q_k} Q_{ik}(\mathbf{q},\lambda). \qquad (2.40)$$

Можно также предположить, что коэффициенты диффузии симметричны:

$$Q_{ik} = Q_{ki}. \qquad (2.41)$$

Определим следующие новые коэффициенты:

а) коэффициенты необратимого дрейфа

$$D_i(\mathbf{q},\lambda) = \frac{1}{2}[K_i(\mathbf{q},\lambda) + \varepsilon_i K_i(\tilde{\mathbf{q}},\tilde{\lambda})] \equiv D_i^{\text{необр}}; \qquad (2.42)$$

б) коэффициенты обратимого дрейфа

$$J_i(\mathbf{q},\lambda) = \frac{1}{2}[K_i(\mathbf{q},\lambda) - \varepsilon_i K_i(\tilde{\mathbf{q}},\tilde{\lambda})] \equiv D_i^{\text{обр}}. \qquad (2.43)$$

При обращении времени J_i преобразуется как q_i.

Запишем стационарное решение уравнения Фоккера–Планка в виде

$$f(\mathbf{q},\lambda) = \mathcal{N} e^{-\Phi(\mathbf{q},\lambda)}, \qquad (2.44)$$

где $\mathcal{N}$ – нормировочная константа, а Ф можно интерпретировать как обобщенный термодинамический потенциал. Необходимые и достаточные условия, которые должны выполняться для того, чтобы не нарушался принцип детального равновесия, сводятся к соотношениям

$$Q_{ik}(\mathbf{q}, \lambda) = \varepsilon_i \varepsilon_k Q_{ik}(\tilde{\mathbf{q}}, \tilde{\lambda}), \tag{2.45}$$

$$D_i - \frac{1}{2}\sum_k \frac{\partial Q_{ik}}{\partial q_k} = -\frac{1}{2}\sum_k K_{ik}\frac{\partial \Phi}{\partial q_k}, \tag{2.46}$$

$$\sum_i \left(\frac{\partial J_i}{\partial q_i} - J_i \frac{\partial \Phi}{\partial q_i}\right) = 0. \tag{2.47}$$

Если у матрицы диффузии существует обратная матрица, то уравнение (2.46) можно разрешить относительно градиента обобщенного термодинамического потенциала Ф:

$$\frac{\partial \Phi}{\partial q_i} = \sum_k (Q^{-1})_{ik}\left(\sum_l \frac{\partial Q_{kl}}{\partial q_l} - 2D_k\right) \equiv A_i. \tag{2.48}$$

Это показывает, что из соотношения (2.48) следует условие интегрируемости

$$\frac{\partial}{\partial q_j} A_i = \frac{\partial}{\partial q_i} A_j. \tag{2.49}$$

Оно же есть условие на коэффициенты дрейфа и диффузии, определяемое правой частью соотношения (2.48). Подставляя A_i и A_j из соотношения (2.48), преобразуем условие (2.47) к виду

$$\sum_i \left[\frac{\partial J_i}{\partial q_i} - J_i \sum_k (Q^{-1})_{ik}\left(\sum_l \frac{\partial Q_{kl}}{\partial q_l} - 2D_k\right)\right] = 0. \tag{2.50}$$

Таким образом, условия, при которых должно выполняться детальное равновесие, определяются в конце концов соотношениями (2.45), (2.49) и (2.50). Соотношение (2.46), или эквивалентное ему условие (2.48), позволяет определить Ф с помощью одних лишь квадратур, т. е. однократного интегрирования. Таким образом, стационарное уравнение Фоккера–Планка может быть определено в явном виде.

2.5. Интегралы по траекториям

Нестационарные решения уравнения Фоккера–Планка могут быть представлены в виде интегралов по траекториям. Для простоты мы ограничимся в дальнейшем случаем, когда коэффициент диффузии Q не зависит от $\mathbf{q}$. Рассмотрим сначала одномерный случай, когда вектор $\mathbf{q}$ сводится к одной переменной. Разделим интервал времени t на эквидистантные шаги

$$t_0, t_1 = t_0 + \tau, \ldots, t_N = t_0 + \tau N. \tag{2.51}$$

Функцию распределения f в момент времени t можно представить в виде кратного интеграла по всем промежуточным положениям q_0, q_1, q_2, ... (см. рис. 2.2). В явном виде интеграл по траекториям можно представить следующим образом:

$$f(q,t) = \lim_{\substack{N \to \infty \\ N\tau = t}} \int \dots \int Dq e^{-G/2} f(q', t_0), \tag{2.52}$$

где мы для краткости ввели обозначения

$$D_q = (2Q\tau\pi)^{-N/2} dq_0, \dots, dq_{N-1}, \tag{2.53}$$

$$G = \sum_{\nu} \frac{\tau \left[\frac{q_\nu - q_{\nu-1}}{\tau} - K(q_{\nu-1}) \right]^2}{Q}. \tag{2.54}$$

Вывод формулы (2.52) дан в моей книге «Синергетика». Рассмотрим теперь обобщение формулы (2.52) на случай многомерного вектора $\mathbf{q}$ с n компонентами. Мы по-прежнему предполагаем, что Q_{mn} не зависит от переменной состояния $\mathbf{q}$. Результат получается следующий:

$$f(\mathbf{q},t) = \lim_{\substack{N \to \infty \\ N\tau = t}} \int \dots \int Dq e^{-G/2} f(\mathbf{q}', t_0), \tag{2.55}$$

где мы воспользовались следующими сокращенными обозначениями:

$$D_q = \prod_{\mu=0}^{N-1} \{(2\pi\tau^{-n/2}(\det Q)^{-1/2}\} (dq_1, \dots, dq_n)_\mu, \tag{2.56}$$

$$\mathbf{q}_N = \mathbf{q}; \ \mathbf{q}_0 = \mathbf{q}', \tag{2.57}$$

$$G = t \sum (\dot{q}_\nu^T - K_{\nu-1}^T) Q^{-1} (\dot{q}_\nu - K_{\nu-1}), \tag{2.58}$$

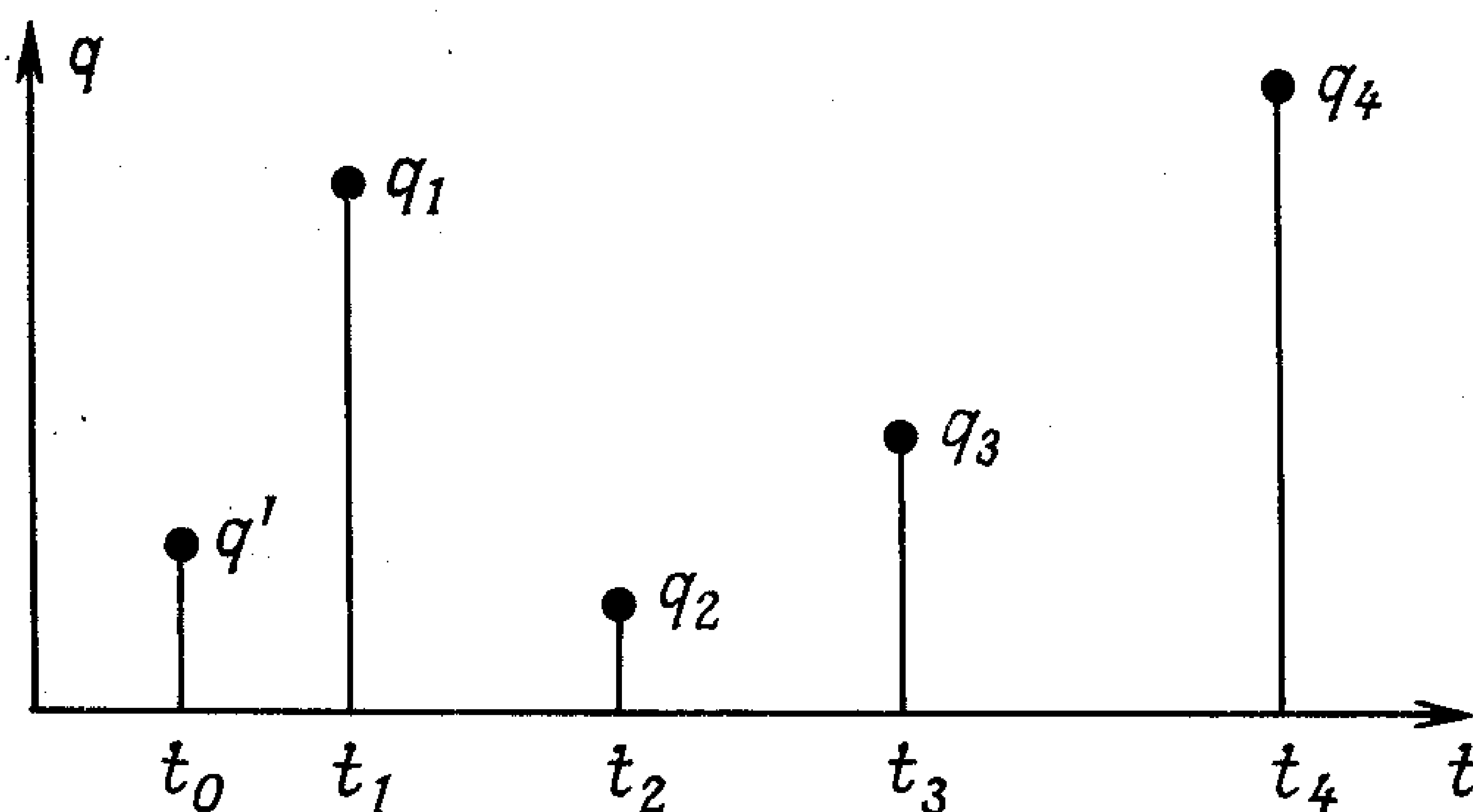


Рис. 2.2. Положения частиц, или состояние системы, в различные моменты времени.

$$\dot{\mathbf{q}}_V = \tau^{-1}(\mathbf{q}_V - \mathbf{q}_{V-1}),\ \mathbf{K}_{V-1} = \mathbf{K}(\mathbf{q}_{V-1}), \qquad (2.59)$$

где T означает транспонированный вектор.

Наконец, напомним читателю об основном кинетическом уравнении. Рассмотрим вектор $\mathbf{m}$ из дискретного пространства состояний. Нас интересует распределение вероятности $P(\mathbf{m},t)$. Если мы имеем дело с марковским процессом, то P удовлетворяет основному кинетическому уравнению

$$\frac{dP(\mathbf{m},t)}{dt} = \sum_{\mathbf{n}} w(\mathbf{m},\mathbf{n})P(\mathbf{n}) - \sum_{\mathbf{n}} w(\mathbf{n},\mathbf{m})\mathbf{P}(\mathbf{m}). \qquad (2.60)$$

Найти явное решение уравнения (2.60) весьма трудно. Но в частном случае, когда в системе достигается детальное равновесие, стационарное распределение вероятности может быть получено в явном виде. Дело в том, что когда устанавливается детальное равновесие, $\mathbf{P}$ удовлетворяет соотношению

$$w(\mathbf{n},\mathbf{m})P(\mathbf{m}) = w(\mathbf{m},\mathbf{n})P(\mathbf{n}). \qquad (2.61)$$

Тогда соответствующее стационарному состоянию решение уравнения (2.60) может быть записано в явной форме как

$$P(\mathbf{m}) = Ne^{\Phi(\mathbf{m})}, \qquad (2.62)$$

где $\Phi(\mathbf{m})$ определяется выражением

$$\Phi(\mathbf{m}) = \Phi(\mathbf{n}_0) + \sum_{j=0}^{N-1} \ln\left\{\frac{w(\mathbf{n}_{j+1},\mathbf{n}_j)}{w(\mathbf{n}_j,\mathbf{n}_{j+1})}\right\}, \qquad (2.63)$$

$$\mathbf{m} \equiv \mathbf{n}_N. \qquad (2.64)$$

2.6. Уменьшение сложности, параметры порядка и принцип подчинения

В этом разделе мы займемся исследованием систем, состоящих из большого числа частей. Нас будут интересовать качественные изменения в поведении системы. Для этого нам понадобится выполнить несколько преобразований переменных и уравнений, которым они удовлетворяют. Затем в разд. 2.7 и 2.8 мы изложим некоторые важные приложения. Начнем с вектора состояния $\mathbf{q}$, описывающего всю систему на микроскопическом или мезоскопическом уровне

$$\mathbf{q} = \mathbf{q}(\mathbf{x},t). \qquad (2.65)$$

В дальнейшем мы будем предполагать, что вектор состояния зависит от пространственной координаты $\mathbf{x}$, поэтому, не уменьшая общности, будем считать, что $\mathbf{q}$ удовлетворяет уравнению типа

$$\dot{\mathbf{q}} = \mathbf{N}(\mathbf{q},\alpha) + \mathbf{F}(t). \qquad (2.66)$$

В этом уравнении $\mathbf{N}$ есть нелинейная функция от $\mathbf{q}$, которая может также содержать дифференциальные операторы (например, оператор Лапласа, дифференцирующий $\mathbf{q}$ по пространственным координатам), α – управляющий параметр, например мощность накачки лазера, и $\mathbf{F}(t)$ – флуктуирующая сила. Разобъем наш последующий анализ на несколько этапов.

2.6.1. Анализ устойчивости по линейному приближению

В дальнейшем мы будем предполагать, что при некотором фиксированном значении управляющего параметра α_0 решение детерминистического уравнения известно, т. е. $\mathbf{q}_0$ удовлетворяет уравнению

$$\dot{\mathbf{q}} = \mathbf{N}(\mathbf{q},\alpha_0). \tag{2.67}$$

Исследуем, как ведет себя решение при изменении управляющего параметра. Для этого предположим, что

$$\alpha: \mathbf{q} = \mathbf{q}_0 + \mathbf{w}. \tag{2.68}$$

Будем считать, что $\mathbf{q}_0$ гладко зависит от α:

$$q_0 = q_0(\alpha). \tag{2.69}$$

Требуется исследовать устойчивость этого решения $\mathbf{q}_0$. Для этого подставим предполагаемую зависимость (2.68) в уравнение (2.66), в котором отброшен член $\mathbf{F}(t)$, и получим

$$\dot{\mathbf{q}}_0 + \dot{\mathbf{w}} = \mathbf{N}(\mathbf{q}_0 + \mathbf{w},\alpha). \tag{2.70}$$

Предполагая, что $\mathbf{w}$ – малая величина, разложим правую часть уравнения (2.70) в степенной ряд по $\mathbf{w}$ и сохраним только два главных члена

$$\dot{\mathbf{q}}_0 + \dot{\mathbf{w}} = \mathbf{N}(\mathbf{q}_0,\alpha) + L(\mathbf{q}_0)\mathbf{w} + \ldots. \tag{2.71}$$

С учетом уравнения (2.67) первый член в левой части уравнения (2.70) взаимно уничтожается с первым членом в правой части уравнения (2.71), после чего остается уравнение

$$\dot{\mathbf{w}} = L(\mathbf{q}_0)\mathbf{w}. \tag{2.72}$$

Заметим, что оператор L, действующий на $\mathbf{q}_0$, может содержать дифференциальные операторы, которые действуют на пространственные координаты величины $\mathbf{w}$. Тем не менее общее решение уравнения (2.72) может быть записано в виде

$$\mathbf{w}(t) = e^{\lambda_t}\mathbf{v}, \tag{2.73}$$

где $\mathbf{v}$ – не зависящий от времени вектор.

Были рассмотрены и более общие случаи, а именно такие, в которых $\mathbf{q}_0$ – периодическая или квазипериодическая функция. За подробным изложением результатов, полученных в этом направлении, я

вынужден отослать читателя к моей книге «Синергетика. Иерархии неустойчивостей». Для дальнейшего важно лишь различать так называемые «неустойчивые» и «устойчивые» моды. Неустойчивыми модами называются такие, для которых

$$\lambda > 0. \tag{2.74}$$

Обозначим их через

$$\lambda_u, \mathbf{v}_u. \tag{2.75}$$

Устойчивыми модами называются такие, для которых

$$\lambda < 0. \tag{2.76}$$

Их мы обозначим через

$$\lambda_s, \mathbf{v}_s. \tag{2.77}$$

Подчеркнем, что термины «устойчивый» и «неустойчивый» относятся только к анализу по линейному приближению. Действительно, оказывается, что так называемые неустойчивые моды, вообще говоря, стабилизируются за счет взаимодействия с устойчивыми модами. Заметим также, что наш подход полностью нелинейный и анализ устойчивости по линейному приближению служит только для того, чтобы найти адекватную систему отсчета для представления искомого решения $\mathbf{q}$ уравнения (2.66).

2.6.2. Преобразование эволюционных уравнений

Чтобы решить уравнение (2.66) в нелинейном и стохастическом случае, примем гипотезу: пусть

$$\mathbf{q} = \mathbf{q}_0 + \sum_u \xi_u(t)\,\mathbf{v}_u + \sum_s \xi_s(t)\,\mathbf{v}_s. \tag{2.78}$$

Если L содержит дифференциальные операторы, действующие на пространственные переменные, то $\mathbf{v}$ зависит от пространственных координат:

$$\mathbf{v} = \mathbf{v}(\mathbf{x}). \tag{2.79}$$

Подставляя формулу (2.78) в уравнение (2.66) и разбивая обе части на составляющие $\mathbf{v}_u$ и $\mathbf{v}_s$, получаем уравнения для амплитуд ξ_u и ξ_s мод:

$$\dot{\xi}_u = \lambda_u \xi_u + N_u(\xi_u, \xi_s) + \mathbf{F}_u(t), \tag{2.80}$$

$$\dot{\xi}_s = \lambda_s \xi_s + N_s(\xi_u, \xi_s) + \mathbf{F}_s(t). \tag{2.81}$$

Индексы u и s используются нами для двух целей. Во-первых, они указывают, с амплитудами каких мод (неустойчивых или устойчивых) мы имеем дело. Во-вторых, они позволяют нумеровать отдельные компоненты векторов ξ_u и ξ_s. Например, мы можем сделать так,

чтобы индексы u и s принимали следующие значения: $u = 1, \ldots, M$ и $s = M + 1, \ldots$. По смыслу станет ясно, как надлежит понимать индексы u и s. Амплитуды ξ_u называются параметрами порядка.

2.6.3. *Принцип подчинения*

Преобразование (2.78) уравнения (2.66) не уменьшает сложности, т.е. уравнения (2.80) и (2.81) полностью эквивалентны уравнению (2.66). Однако синергетический принцип подчинения позволяет исключить из уравнений (2.80) и (2.81) амплитуды подчиненных мод с помощью явной формулы

$$\xi_s(t) = f_s[\xi_u(t), t]. \tag{2.82}$$

Явное построение функции f_s описано в моей книге «Синергетика. Иерархии неустойчивостей», а в частных случаях и в моей первой книге «Синергетика». Поэтому сейчас мы лишь проиллюстрируем содержание формулы (2.82) на простом примере, позволяющем показать самую суть принципа подчинения. Для этого рассмотрим уравнения для амплитуд ξ_u и ξ_s, а именно уравнения (2.80) и (2.81), которые запишем в следующем виде:

$$\dot{\xi}_u = \lambda_u \xi_u + h_u(\xi_u, \xi_s) + F_u(t), \tag{2.83а}$$

$$\dot{\xi}_s = \lambda_s \xi_s + g_s(\xi_u) + q_s k_s(\xi_u) + F_s(t). \tag{2.83б}$$

Предполагается, что h_u – нелинейная функция, разложение которой в ряд Тейлора начинается с членов не ниже второго порядка по ξ_u. Аналогично g_s – функция, разложение которой в ряд Тейлора начинается с той же степени, что и разложение функции h_u. И можно показать, что разложение ξ_s начинается с членов не ниже второго порядка по ξ_u. В своей простейшей форме принцип подчинения сводится к подстановке в уравнение (2.83б) $\dot{\xi}_s = 0$.

Удерживая только главные члены, получаем

$$\xi_s \approx -\frac{1}{\lambda_s} g_s(\xi_u) - \frac{1}{\lambda_s} F_s(t). \tag{2.84}$$

Этот результат может быть строго доказан для низшего порядка в F_s и ξ_u. Исследуем теперь, что означает принцип подчинения для решения уравнения Фоккера–Планка. Для этого преобразуем уравнение Фоккера–Планка, перейдем от старого вектора состояния $\mathbf{q}$ к новым переменным ξ_u, ξ_s:

$$\mathbf{q} \to \xi_u, \xi_s. \tag{2.85}$$

Уравнение Фоккера–Планка примет при этом общий вид

$$\dot{P}(\xi_u, \xi_s; t) = L(\xi_u, \xi_s)\mathbf{P}(\xi_u, \xi_s; t), \tag{2.86}$$

где L – линейный оператор. Рассмотрим стационарное решение урав-

нения (2.86). Его всегда можно представить в виде

$$P(\xi_u, \xi_s) = P(\xi_s \mid \xi_u) f(\xi_u), \tag{2.87}$$

где слева стоит совместная вероятность, в то время как P в правой части – условная вероятность. Что же касается f, то это – функция распределения только параметров порядка. Принцип подчинения (в своем главном приближении) означает, что условную вероятность P в правой части уравнения (2.87) можно записать в более специальном виде как

$$P(\xi_s \mid \xi_u) = \prod_s P_s(\xi_s \mid \xi_u). \tag{2.88}$$

Нашим результатом (2.84) можно воспользоваться для того, чтобы на явном примере показать, как может выглядеть разложение (2.88) в этом приближении низшего порядка для принципа подчинения. Флуктуирующие силы имеют, как обычно, гауссовское распределение, т.е. вероятность найти F_s в интервале $F \to F + dF$ определяется выражением

$$P(F \leqslant F_s \leqslant F + dF) = N'_{\text{exp}}(-F_s^2/Q')\,dF. \tag{2.89}$$

Разрешим соотношение (2.84) относительно F_s:

$$F_s = -\lambda_s\left[\xi_s + \frac{1}{\lambda_s} g_s(\xi_u)\right]. \tag{2.90}$$

Это позволяет нам определить вероятность с помощью формулы (2.89). В результате получаем

$$P(\xi_s \mid \xi_u) = N_{\text{exp}}\left\{-\left[\xi_s + \frac{1}{\lambda_s} g_s(\xi_u)\right]^2 \Big/ Q\right\} d\xi_s, \tag{2.91}$$

где мы для краткости положили

$$Q^{-1} = Q'^{-1}\lambda_s^2, \tag{2.92}$$

$$N = N'\lambda_s. \tag{2.93}$$

2.7. Неравновесные фазовые переходы

Во многих случаях, представляющих практический интерес, число параметров порядка может быть очень мало или даже равно единице, в то время как число подчиненных мод по-прежнему очень велико. Рассмотрим случай одного параметра порядка и условимся для простоты отбрасывать индекс u:

$$\xi_u \to \xi,\ F_u \to F,\ \lambda_u \to \lambda. \tag{2.94}$$

Тогда типичное уравнение для параметра порядка будет иметь вид

$$\dot{\xi} = \lambda\xi - \beta\xi^3 + F(t), \tag{2.95}$$

как это показано в синергетике.

Флуктуирующая сила $F(t)$ удовлетворяет соотношению

$$\langle F(t)F(t')\rangle = Q\delta(t - t'). \tag{2.96}$$

Если, кроме того, предположить, что F имеет гауссовское распределение, то мы получим соответствующее уравнению (2.95) уравнение Фоккера–Планка в виде

$$\dot{f}(\xi; t) = -\frac{\partial}{\partial\xi}[(\lambda\xi - \beta\xi^3)f] + \frac{Q}{2}\frac{\partial^2}{\partial\xi^2}f. \tag{2.97}$$

Рассмотрим стационарное состояние, для которого

$$\dot{f} = 0. \tag{2.98}$$

При этом условии уравнение (2.97) легко интегрируется:

$$f = N_{\exp}\left[Q^{-1}\left(\lambda\xi^2 - \frac{1}{2}\beta\xi^4\right)\right]. \tag{2.99}$$

Формула (2.99) – явный и весьма характерный пример функции распределения f, удовлетворяющей соотношению (2.87). Функции распределения для параметров порядка могут быть найдены в явном виде и в случае нескольких параметров порядка, если выполняется принцип детального равновесия. Предположим, что λ_u – вещественная величина. Частный случай принципа детального равновесия состоит в следующем. Уравнение Ланжевена имеет вид

$$\dot{\xi}_u = \lambda_u\xi_u - \frac{\partial V}{\partial\xi_u} + F_u(t), \tag{2.100}$$

где V – нелинейная функция от ξ_u. Предполагается, что

$$\langle F_u(t)F_{u'}(t')\rangle = \delta_{uu'}Q\delta(t - T). \tag{2.101}$$

Тогда решение уравнения Фоккера–Планка, соответствующего уравнению Ланжевена (2.100), можно записать в общем виде

$$f(\xi_u) = N_{\exp}\left[-\frac{2V(\xi_u)}{Q}\right]. \tag{2.102}$$

Даже если принцип детального равновесия не выполняется, то функция распределения односвязна; ее всегда можно представить в виде

$$f(\xi_u) = N_{\exp}[-\Phi(\xi_u)], \tag{2.103}$$

где Φ играет роль обобщенного термодинамического потенциала. Предположим теперь, что при изменении управляющего параметра происходит мягкий переход, в результате которого система становится неустойчивой и входит в новую область. В этом случае мы

можем предположить, что параметры порядка по-прежнему малы, например, что

$$\xi_u \sim \lambda_u^{\varkappa}, \quad \varkappa > 0. \tag{2.104}$$

В этом случае мы можем разложить Φ в степенной ряд по ξ_u:

$$\Phi(\xi_u) = \sum_u \lambda_u \xi_u^2 + \sum_{u,u'} c_{uu'} \xi_u \xi_{u'} + \sum_{u,u',u''} c_{uu'u''} \xi_u \xi_{u'} \xi_{u''} + \\ + \sum_{u,u',u'',u'''} \xi_u \xi_{u'} \xi_{u''} \xi_{u'''}. \tag{2.105}$$

Для большинства неравновесных фазовых переходов достаточно удерживать первые четыре степени, хотя в исключительных случаях необходимо удерживать и более высокие степени. Вообще говоря, системы обладают внутренними симметриями, которые приводят к определенным соотношениям между коэффициентами c членов каждого порядка.

2.8. Образование структур

Теперь мы хотим показать, каким образом описанный выше формализм может описывать образование структур. Если мы рассматриваем непрерывно распределенные среды, характеризуемые пространственной координатой $\mathbf{x}$, то оператор L в линеаризации (2.72), вообще говоря, содержит производные по пространственной координате. В этом случае $\mathbf{w}$ и тем самым величина $\mathbf{v}$, входящая в уравнение (2.73), становятся функциями пространственной координаты $\mathbf{x}$ (соотношение (2.79)).

В совершенно общем случае решение $\mathbf{q}$ нелинейного уравнения (2.66) может быть записано в виде

$$\mathbf{q} = \mathbf{q}_0 + \sum_u \xi_u(t)\, \mathbf{v}_u(\mathbf{x}) + \sum_s \xi_s(t)\, \mathbf{v}_s(\mathbf{x}), \tag{2.106}$$

таком же, как в нашей предыдущей формуле (2.78).

Оказывается, что ξ_u в общем случае на порядок величины больше, чем ξ_s, т. е. развивающаяся во времени структура определяется главным образом первой суммой (по u) в разложении (2.106), которую мы назовем остовом моды. Если имеется только один параметр порядка и $\mathbf{v}_u$ имеет вид

$$\mathbf{v}_u = L^{-1/2} \sin kx, \tag{2.107}$$

тогда разложение (2.106) по существу вырождается в более краткое выражение

$$\mathbf{q} = \mathbf{q}_0 + \xi_u(t)\, L^{-1/2} \sin kx. \tag{2.108}$$

Во многих случаях ξ_u удовлетворяет уравнению (2.95), которое описывает рост ξ_u от начальной флуктуации до окончательных размеров (рис. 2.3, 2.4). Совершенно очевидно, что при большем числе параметров порядка и/или более сложных функциях v_u могут возникать гораздо более сложные структуры, чем те, которые описываются формулами (2.107) и (2.108). Таким образом, развитая выше теория способна приводить к заключению и о возникновении пространственных структур в сложных системах. Наш пример – не более, чем краткое напоминание того, что было подробно изложено в моих книгах «Синергетика» и «Синергетика. Иерархии неустойчивостей». Я хочу лишь подчеркнуть, что в этих книгах были рассмотрены существенно более общие случаи, такие, как структуры, развивающиеся во времени, а именно периодические циклы, квазипериодическое движение или хаос. Но для последующего изложения этого краткого напоминания в большинстве случаев оказывается вполне достаточно.

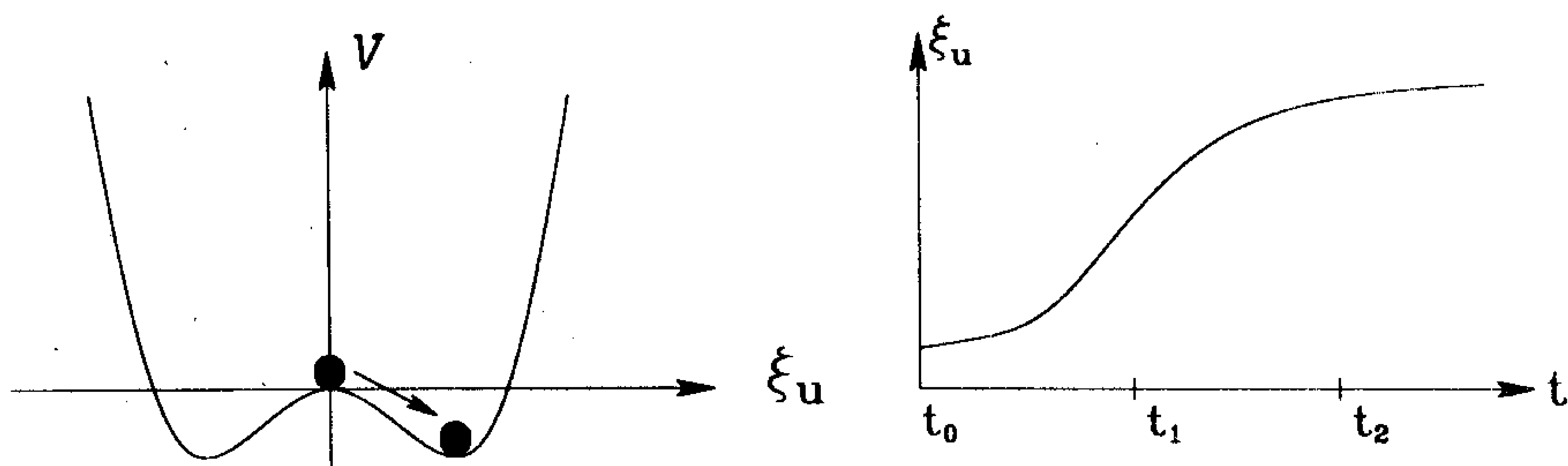


Рис. 2.3. Пример зависимости параметра порядка ξ_u от времени. Слева: потенциал V, в котором может двигаться фиктивная частица с координатой ξ_u. Справа: временная зависимость параметра порядка ξ_u.

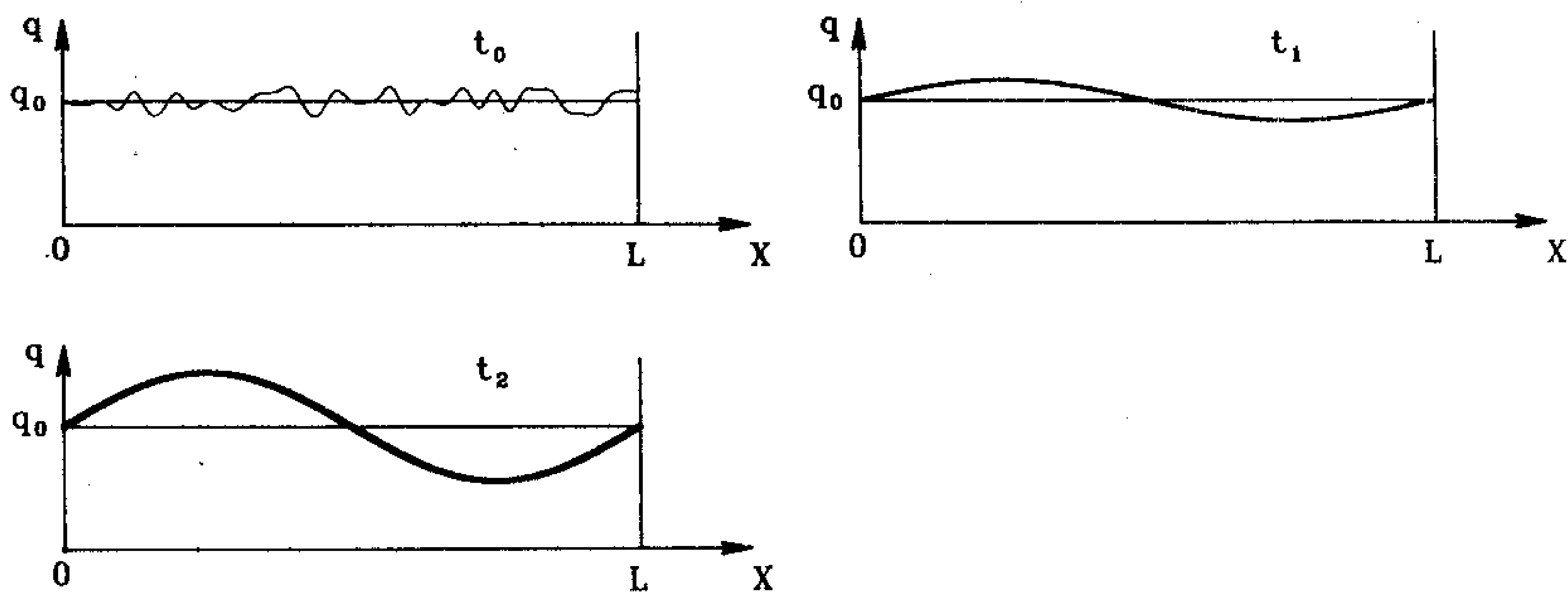


Рис. 2.4. Эволюция пространственной структуры, описываемой уравнением (2.108) во времени. Параметр порядка изменяется так, как показано на рис. 2.3.

Глава 3

... и снова назад: принцип максимума информации (ПМИ)

3.1. Некоторые основные идеи

В этой главе мы пытаемся ответить на следующий вопрос. Пусть заданы *макроскопические величины* рассматриваемой нами системы. Требуется разработать процедуру, которая позволила бы нам вывести распределение вероятности для *макроскопических* или даже *микроскопических* переменных. Иначе говоря, мы начинаем с макроскопического мира и хотим прийти к тем или иным заключениям относительно микроскопического мира. В зависимости от рода рассматриваемой системы адекватные макроскопические величины могут быть совершенно различными. В закрытых физических системах, к которым приложима термодинамика, такие величины – энергия, число частиц и т. д.; общую процедуру мы проиллюстрируем этим примером в гл. 4. В открытых системах, например в физике и биологии, адекватными макроскопическими величинами могут оказаться, например, интенсивности и флуктуации интенсивностей. Главная цель последующих глав этой книги как раз и будет состоять в анализе *открытых систем.*

Так как исходным пунктом нашего подхода является понятие информации, мы обсудим его в разд. 3.1.

С помощью своего рода новой интерпретации теории вероятностей мы получаем возможность проникнуть в существо того, чем занимается, казалось бы, совершенно другая дисциплина, а именно теория информации. Последовательность бросаний монеты с исходами 0 и 1. Условимся затем интерпретировать 0 и 1 как тире и точку в азбуке Морзе. Как мы все знаем, с помощью азбуки Морзе можно передавать сообщения, поэтому ничто не мешает нам приписать какой-нибудь последовательности символов вполне определенный смысл. Иначе говоря, последовательность символов может быть носителем информации. В теории информации мы пытаемся найти меру количества информации.

Рассмотрим простой пример. Пусть имеется R_0 различных возможных событий («реализаций»), которые априори равновероятны. Например, при бросании монеты мы должны иметь события 1 и 0 и $R_0 = 2$. В случае игральной кости число возможных исходов равно 6, поэтому $R_0 = 6$. Таким образом, исход бросания монеты или игральной кости интерпретируется как рецепт получения сообщения, и реализуется только один из возможных исходов R_0. Ясно, что чем

больше R_0, тем выше неопределенность до получения сообщения и тем большее количество информации приобретается после получения сообщения. Следовательно, всю процедуру можно рассматривать так: в самом начале мы не располагали никакой информацией I_0, т.е. при R_0 равновероятных исходах $I_0 = 0$.

В самом конце мы располагаем информацией $I_1 \neq 0$ при $R_1 = 1$, т.е. при одном исходе. Предположим, что мы хотим ввести меру количества информации I, которое явно должно быть связано с R_0. Чтобы получить представление о том, какая связь должна возникнуть между R_0 и I, потребуем аддитивности I для независимых событий. Таким образом, если мы имеем два независимых множества с R_{01} или R_{02}, то общее число исходов равно

$$R_0 = R_{01} R_{02} \tag{3.1}$$

и мы требуем, чтобы

$$I(R_{01} R_{02}) = I(R_{01}) + I(R_{02}). \tag{3.2}$$

Этому требованию можно удовлетворить, если выбрать

$$I = K \ln R_0, \tag{3.3}$$

где K – константа. Можно даже показать, что (3.3) – единственное решение функционального уравнения (3.2). Константа K пока произвольна и может быть зафиксирована с помощью какого-нибудь требования. Обычно используется следующее требование. Рассмотрим так называемую двоичную систему, в которой имеются только два символа (две буквы). Это могут быть «орел» и «решка» при бросании монеты, ответы типа «да» и «нет» или числа 0 и 1 в двоичной системе. Построив все возможные «слова» (или последовательности) длины n, мы получим $R = 2^n$ реализаций. Мы хотим идентифицировать количество информации I с числом n в такой системе. Это означает, что мы требуем выполнения равенства

$$I = K \ln R \equiv K \ln 2 = n, \tag{3.4}$$

которому соответствует выбор

$$K = \frac{1}{\ln^2} = \log_2 e. \tag{3.5}$$

Другая запись соотношения (3.4) при том же выборе K имеет вид

$$I = \log_2 R. \tag{3.4a}$$

Так как одна позиция в последовательности символов (знаков) двоичной системы называется битом, информация при таком определении измеряется прямо в битах. Например, если $R = 8 = 2^3$, то мы получаем информацию $I = 3$ бита, а в общем случае при $R = 2^n$ – информацию $I = n$ битов. Определение информации (3.3) легко обобщается на случай, когда мы имеем первоначально R_0 равновероятных

выборов, а в конце – R_1 равновероятных выборов. Информация при этом оказывается равной

$$I = K \ln R_0 - K \ln R_1. \qquad (3.6)$$

Это выражение сводится к предыдущему (формуле (3.3)) при $R_1 = 1$. Простым примером может служить игральная кость. Условимся считать, что выпадение четного числа очков означает выигрыш, а выпадение нечетного числа очков – проигрыш. Тогда $R_0 = 6$ и $R_1 = 3$. В этом случае количество информации оказывается таким же, как в случае бросания монеты с двумя возможными исходами бросания.

Выведем теперь более удобное выражение для информации. Рассмотрим сначала следующий пример упрощенной азбуки Морзе, состоящей лишь из двух символов – точки и тире (в настоящей азбуке Морзе имеются три символа: точка, тире и пробел). Пусть слово длины G содержит N_1 тире и N_2 точек, где

$$N_1 + N_2 = N. \qquad (3.7)$$

Спрашивается, какую информацию мы получим, приняв такое слово. Действуя в духе теории информации, мы должны вычислить общее число слов, которые могут быть образованы из этих двух символов при фиксированных N_1, N_2. Анализ очень прост. Подсчитывая, сколькими способами можно расставить точки и тире по N местам, получаем, что число возможностей равно

$$R = \frac{N!}{N_1!\, N_2!}. \qquad (3.8)$$

Иначе говоря, R – число сообщений, которые могут быть переданы с помощью N_1 тире и N_2 точек. Теперь мы хотим установить, какое количество информации приходится на один символ, т.е. оценить величину $i = I/N$. Подставляя выражение (3.8) в формулу (3.3), получаем

$$I = K \ln R = K [\ln N! - \ln N_1! - \ln N_2!]. \qquad (3.9)$$

Воспользуемся аппроксимацией

$$\ln N! \approx N(\ln N - 1) \qquad (3.10)$$

формулы Стирлинга, которая обладает достаточной точностью при $N > 100$. Тогда

$$I \approx K\, [N(\ln N - 1) - N_1(\ln N_1 - 1) - N_2(\ln N_2 - 1)], \qquad (3.11)$$

и, принимая во внимание соотношение (3.7), получаем

$$i \equiv \frac{I}{N} \approx -K \left[\frac{N_1}{n} \ln \frac{N_1}{N} + \frac{N_2}{N} \ln \frac{N_2}{N} \right]. \qquad (3.12)$$

Введем теперь величину, которую можно интерпретировать как вероятность найти знак «тире» или «точка». Эта вероятность равна

относительной частоте, с которой встречается тире или точка

$$p_j = \frac{N_j}{N}, \quad j = 1,2. \tag{3.13}$$

С учетом этого наша окончательная формула принимает следующий вид:

$$i = \frac{I}{N} = -K(p_1 \ln p_1 + p_2 \ln p_2). \tag{3.14}$$

Выражение (3.14) легко обобщается на случай, когда алфавит содержит не два символа, а больше. Рассуждая аналогичным образом, мы заключаем, что в общем случае информация на один символ определяется выражением

$$i = -K \sum_j p_j \ln p_j, \tag{3.15}$$

где p_j – относительная частота, с которой встречаются символы. Из этой интерпретации ясно, что величина i может служить количественной характеристикой при передаче информации и т. д.

Прежде чем мы продолжим, необходимо сказать несколько слов о том, в каком смысле следует понимать здесь слово «информация». Как известно, такие характеристики информации, как «полезная» или «бесполезная», «имеющая смысл» или «не имеющая смысла» в теорию информации не входят. Например, в случае азбуки Морзе, о которой мы говорили выше, очень многие слова лишены смысла. Информация в том смысле, о котором идет речь в нашей книге, скорее характеризует, насколько редко происходят события. Хотя это существенно ограничивает теорию, она все же оказывается необычайно полезной.

Полученное нами выражение для информации можно рассматривать с двух совершенно различных точек зрения. С одной стороны, мы можем предположить, что вероятности p_i заданы своими численными значениями, и по формуле (3.3) получить соответствующее этим значениям число для I. Однако гораздо важнее вторая точка зрения, согласно которой информацию I надлежит рассматривать как функцию от p_i, так что если значения вероятностей p_i изменяются, то значение I изменяется в соответствии с ними. Чтобы пояснить такую интерпретацию информации, мы кратко упомянем об одном приложении, которое позднее рассмотрим гораздо более подробно. Предположим, что имеется газ, состоящий из атомов, которые свободно движутся в сосуде. Нас будет интересовать пространственное распределение атомов газа. Разделим сосуд на M ячеек одинаковой величины и обозначим через N_k число частиц в k-й ячейке. Общее число частиц равно N. Относительная частота обнаружения частицы в k-й ячейке равна отношению

$$\frac{N_k}{N} = p_k, \quad k = 1, 2, ..., M. \tag{3.16}$$

Величину p_k можно рассматривать как функцию распределения частиц по ячейкам k. Так как все ячейки имеют одинаковые размеры и не отличаются по своим физическим свойствам, мы ожидаем, что частицы с равной вероятностью будут находиться в любой ячейке, т. е. что

$$p_k = \frac{1}{M}. \tag{3.17}$$

Выведем теперь формулу (3.17) из свойств информации. Информация может заключаться в следующем. До того как мы производим измерение или получаем сообщение, существуют R возможных исходов, или, иначе говоря, $K \ln R$ есть мера нашего незнания. Другая точка зрения на ту же ситуацию состоит в том, что R задает число реализаций, которые возможны в принципе.

Рассмотрим теперь ансамбль, состоящий из C сосудов, каждый из которых содержит по N атомов газа. Предполагается, что в каждом сосуде частицы распределены в соответствии со своей, отличной от других функцией распределения p_k, т. е. мы имеем набор вероятностей

$$p_k^{(1)},\ p_k^{(2)},\ p_k^{(3)},\ \ldots.$$

Соответственно мы получаем различное число реализаций, т. е. различную информацию. Например, если $N_1 = N$, $N_2 = N_3 = ... = 0$, то $p_1^{(1)} = 1$, $p_2^{(1)} = p_3^{(1)} = ... = 0$, и поэтому $I^{(1)} = 0$. С другой стороны, если $N_1 = N_2 = N_3 = ... = N/M$, то $p_1^{(2)} = 1/M$, $p_2^{(2)} = 1/M, ...$, и тогда $I^{(2)} = -M \log_2(1/M) = M \log_2 M$. Если число сосудов велико, то $I^{(2)}$ – очень большое число.

Таким образом, какой бы сосуд с атомами газа мы ни выбрали, вероятность того, что нам попадется сосуд со второй функцией распределения, гораздо выше, чем вероятность того, что нам попадется сосуд с первой функцией распределения. Это означает, что с наибольшей вероятностью реализуется та функция распределения вероятности p_k, которая соответствует наибольшему числу возможностей R и тем самым несет наибольшую информацию. Следовательно, мы должны потребовать, чтобы условие

$$-\sum p_i \ln p_i = \text{Extr} \tag{3.18}$$

выполнялось при ограничении

$$\sum_{i=1}^{M} p_i = 1 \tag{3.19}$$

(полная сумма вероятностей p_i равна единице). Как станет ясно из дальнейшего, этот принцип имеет фундаментальное значение для приложений к реалистическим системам в физике, химии и биологии, и мы еще вернемся к нему.

Экстремум (3.18) при ограничении (3.19) может быть найден с помощью метода множителей Лагранжа. Суть метода состоит в том, что мы умножаем соотношение (3.19) на пока еще не известный параметр и прибавляем полученное произведение к левой части соотношения (3.18), после чего находим экстремум получившегося выражения. При этом в новом выражении мы можем менять вероятности p_i независимо друг от друга, не учитывая более ограничение (3.19). Проварьировать левую часть выражения

$$-\sum p_i \ln p_i + \lambda \sum p_i = \mathrm{Extr} \tag{3.20}$$

означает взять производную от нее по p_i, что приводит к соотношению

$$-\ln p_i - 1 + \lambda = 0. \tag{3.21}$$

Уравнение (3.21) допускает решение

$$p_i = \mathrm{e}^{\lambda - 1}, \tag{3.22}$$

которое не зависит от индекса i, т. е. все p_i равны. Подставляя их в условие (3.19), мы легко можем определить λ так, чтобы выполнялось равенство

$$M\mathrm{e}^{\lambda - 1} = 1, \tag{3.23}$$

или, иначе говоря, находим, что

$$p_i = \frac{1}{M}, \tag{3.24}$$

как и следовало ожидать в соответствии с формулой (3.17).

3.2. Приращение информации

Выражение (3.15) для информации можно интерпретировать как среднее по f_j

$$i = \sum p_j f_j, \tag{3.25}$$

где

$$f_j = -K \ln p_j, \quad p_j \neq 0 \tag{3.26}$$

и вес равен p_j.

Это наводит на мысль о том, что величину f_j, задаваемую соотношением (3.26), можно интерпретировать как информационное содержание символа с индексом j и величиной p_j, интерпретируемой как вероятность (или относительная частота).

Предположим теперь, что множество измерений привело нас к относительной частоте p_j на символ с индексом j. Предположим теперь, что (возможно, при других условиях) мы определили новую относительную частоту p'_j. Чему равно соответствующее изменение информации Δ_j?

Принимая интерпретацию (3.26), мы сразу же приходим к необходимости определить Δ_j соотношением

$$\Delta_j = K \ln p'_j - K \ln p_j. \tag{3.27}$$

Чтобы найти среднее изменение информации, усредним соотношение (3.27) по новой функции распределения («относительной частоте») p'_j. При этом мы получаем так называемое приращение информации (или «информацию Кульбака»):

$$K(p', p) = \sum_j p'_j \Delta_j = K \sum_j p'_j \ln \frac{p'_j}{p_j}, \tag{3.28}$$

где, разумеется, выполняются условия

$$\sum_j p_j = 1, \tag{3.29}$$

$$\sum_j p'_j = 1. \tag{3.30}$$

Приращение информации $K(p', p)$ обладает следующим важным свойством:

$$K(p', p) \geqslant 0. \tag{3.31}$$

Знак равенства имеет место в том и только в том случае, если

$$p' \equiv p, \text{ т. е. } p'_k = p_k \text{ при всех } k.$$

3.3. Информационная энтропия и ограничения, налагаемые на систему

В этом разделе и в следующей главе нас будут интересовать приложения понятия информации в физике, и мы будем придерживаться традиции обозначать информацию через S и отождествлять константу K в соотношении (3.3) с постоянной Больцмана k_B. По причинам, которые станут ясными в дальнейшем, величину S принято называть информационной энтропией. Поскольку и химические, и биологические системы можно считать физическими системами, наши соображения применимы и к ним. Более того, развиваемый в этой главе общий формализм применим и к другим наукам и областям человеческой деятельности, таким, как обработка информации и т. д. Начнем с основного выражения

$$S = -k_B \sum_j p_i \ln p_i. \tag{3.32}$$

Можно считать, что индекс i описывает индивидуальные особенности частиц или подсистем. Поясним это замечание несколько подробнее. Индекс i может описывать, например, положение частицы газа или ее скорость, или ее положение и скорость. В наших предыдущих примерах индекс i относился к сосудам, заполненным шарами. При более общей интерпретации индекс i соответствует значениям, которые может принимать случайная переменная. В этом разделе мы будем для простоты предполагать, что i – дискретный индекс.

Главная задача, решению которой посвящена эта книга, состоит в том, чтобы найти способы, позволяющие определять p_i (для сравнения вспомним о сосуде с молекулами газа, положение которых требовалось определить). Задача, с которой мы сталкиваемся во многих областях науки, состоит в получении несмещенных оценок, приводящих к таким значениям p_i, которые согласуются со всеми известными сведениями о системе. Рассмотрим идеальный одномерный газ. В этом случае мы заведомо можем измерить, например, положение центра масс. Ограничение («связь») имеет вид

$$\sum_j p_i q_i = M, \tag{3.33}$$

где q_i указывает положение i-й ячейки. Величина M фиксирована и равна Q/N, где Q – координата центра масс, а N – число частиц. Разумеется, существует очень много наборов величин p_i, удовлетворяющих соотношению (3.33). Это позволяет нам выбирать набор $\{p_i\}$ весьма произвольно, т. е. мы можем отдавать одному набору предпочтение перед другим. Как и в обыденной жизни, такая свобода приводит к смещенным оценкам. Но насколько смещены такие оценки? Возвращаясь к примеру с атомами газа, мы можем воспользоваться принципом, о котором шла речь в разд. 3.1: с подавляющей вероятностью реализовываться будут такие распределения, для которых величина (3.32) достигает максимума. Но в силу ограничения (3.33) принимать во внимание следует не все распределения. Если у нас имеется не одно, а несколько ограничений, то наш принцип допускает обобщение. Например, пусть переменная i характеризует различные скорости. Тогда в качестве ограничения мы можем принять условие постоянства полной кинетической энергии $E_{\text{кин}}^{\text{полн}}$ частиц. Если кинетическую энергию частицы с массой m и скоростью v_i обозначить через f_i $[f_i = (m/2)v_i^2]$, то средняя кинетическая энергия, приходящаяся на одну частицу, определяется соотношением

$$\sum_i p_i f_i = E_{\text{кин}}. \tag{3.33a}$$

В общем случае одну i-ю систему можно охарактеризовать величинами $f_i^{(k)}$, $k = 1, 2, \ldots, M$ (которые могут быть пространственными координатами, кинетической энергией или какими-нибудь другими характерными свойствами). Если же величины аддитивны и соответствующие суммы должны оставаться постоянными и равными f_k, то ограничения принимают следующий вид:

$$\sum_i p_i f_i^{(k)} = f_k. \tag{3.34}$$

Кроме того, необходимо добавить обычное ограничение нормировки распределения вероятности

$$\sum_i p_i = 1. \tag{3.35}$$

Задачу нахождения экстремума выражения (3.32) при ограничениях (3.34) и (3.35) можно решить методом множителей Лагранжа λ_k, $k = 1, 2, \ldots, M$ (см. разд. 3.1). Умножим левую часть ограничения (3.34) на λ_k, а левую часть ограничения (3.35) – на $(\lambda - 1)$ и образуем сумму полученных произведений. Вычтем из этой суммы величину $(1/k_B)S$. Коэффициент $1/k_B$ соответствует определенной нормировке множителей λ и λ_k. Проварьируем образовавшуюся сумму по величинам p_i:

$$\delta\left[\frac{1}{k_B}S - (\lambda - 1)\sum_i p_i - \sum_k \lambda_k \sum_i p_i f_i^{(k)}\right] = 0. \tag{3.36}$$

Дифференцируя по p_i и приравнивая полученное выражение нулю, получаем

$$-\ln p_i - 1 - (\lambda - 1) - \sum_k \lambda_k f_i^{(k)} = 0. \tag{3.37}$$

Это соотношение может быть легко разрешено относительно p_i:

$$p_i = \exp\left(-\lambda - \sum_k \lambda_k f_i^{(k)}\right). \tag{3.38}$$

Подставляя выражение (3.38) в условие (3.35), приходим к равенству

$$e^{-\lambda}\sum_i \exp\left(-\sum_k \lambda_k f_i^{(k)}\right) = 1. \tag{3.39}$$

Сумму по i в формуле (3.39) удобно обозначить для краткости особым символом:

$$\sum_i \exp\left(-\sum_k \lambda_k f_i^{(k)}\right) = Z(\lambda_1, \ldots, \lambda_M). \tag{3.40}$$

Мы будем называть величину $Z(\lambda_1, \ldots, \lambda_M)$ функцией распределения. Подставляя выражение (3.40) в равенство (3.39), получаем

$$e^{\lambda} = Z, \tag{3.41}$$

где

$$\lambda = \ln Z. \tag{3.42}$$

Величина λ известна, если величины λ_k определены. Чтобы получить уравнения для λ_k, подставим выражение (3.38) в уравнения (3.34), задающие ограничения, и получим

$$\langle f_i^{(k)} \rangle = \sum_i p_i f_i^{(k)} = e^{-\lambda} \sum_i \exp\left(-\sum_l \lambda_l f_i^{(l)}\right) f_i^{(k)}. \tag{3.43}$$

По своей структуре соотношение (3.43) аналогично соотношению (3.40). Различие между ними состоит в том, что в соотношение (3.43) экспонента входит с множителями $f_i^{(k)}$. Тем не менее сумму, входящую в соотношение (3.43), нетрудно получить из соотношения (3.40), дифференцуруя его по λ_k. Выражая первый множитель в соотношении (3.43) через Z (по формуле (3.41)), получаем

$$\langle f_i^{(k)} \rangle = \frac{1}{Z}\left(-\frac{\partial}{\partial \lambda_k}\right) \sum_i \exp\left(-\sum_i \lambda_l f_j^{(l)}\right), \tag{3.44}$$

или в еще более краткой форме

$$f_k \equiv \langle f_i^{(k)} \rangle = -\frac{\partial \ln Z}{\partial \lambda_k}. \tag{3.45}$$

Так как величины f_k в левой части последнего равенства заданы [см. соотношение (3.34)], а величина Z определяется формулой (3.40), т. е. особым образом зависит от λ_k, соотношения (3.45) можно рассматривать как точную форму системы уравнений для λ_k.

Приведем также формулу, которая понадобится нам в дальнейшем. Подставляя выражение (3.38) в основное соотношение (3.32), получаем

$$\frac{1}{k_B} S_{\max} = \lambda \sum_i p_i - \sum_k \lambda_k \sum_l p_i f_i^{(k)}. \tag{3.46}$$

Формулы (3.34) и (3.35) позволяют преобразовать равенство (3.46) к виду

$$\frac{1}{k_B} S_{\max} = \lambda + \sum_k \lambda_k f_k. \tag{3.47}$$

Таким образом, максимум информационной энтропии может быть выражен через средние значения f_k и множители Лагранжа λ_k. Читатели, знакомые с уравнениями Лагранжа первого рода в механике, вспомнят, что множители Лагранжа имеют там физический смысл сил. Аналогичным образом мы увидим в дальнейшем, что множители Лагранжа λ_k допускают физическую (химическую, биологическую и т. д.) интерпретацию. Выводом формул [(3.38), (3.42), а также (3.32),

(3.45), и (3.47)] мы выполнили нашу первоначальную задачу – найти p и S_{max}.

Выведем теперь еще несколько полезных соотношений. Прежде всего выясним, как изменяется информация S_{max}, если мы изменяем в условии (3.34) функции $f_i^{(k)}$ и f_k. Так как из соотношения (3.47) следует, что S зависит не только от f, но и от λ и λ_k, которые в свою очередь зависят от f, при дифференцировании по f необходимо соблюдать определенную осторожность. Вычислим поэтому сначала изменение λ (см. соотношение (3.42)):

$$\delta\lambda \equiv \delta \ln Z = \frac{1}{Z}\delta Z.$$

Подставляя вместо Z левую часть равенства (3.40), получаем

$$\delta\lambda = e^{-\lambda}\sum_i \sum_k \{-\delta\lambda_k f_i^{(k)} - \lambda_k \delta f_i^{(k)}\} \exp\left(-\sum_l \lambda_l f_i^{(l)}\right).$$

Используя определение (3.38) величин p_i, преобразуем последнее выражение следующим образом:

$$\delta\lambda = -\sum_k [\delta\lambda_k \sum_i p_i f_i^{(k)} + \lambda_k \sum_i p_i \delta f_i^{(k)}].$$

Формула (3.43) и аналогичное выражение для $\langle \delta f_i^{(k)} \rangle$ позволяет записать правую часть последней формулы в виде

$$-\sum_k [\delta\lambda_k \langle f_i^{(k)} \rangle + \lambda_k \langle \delta f_i^{(k)} \rangle]. \tag{3.48}$$

Подставляя выражение (3.48) в формулу (3.47) для δS_{max}, мы замечаем, что вариации величин λ_k выпадают и остается соотношение

$$\delta S_{max} = k_B \sum_k \lambda_k [\delta \langle f_i^{(k)} \rangle - \langle \delta f_i^{(k)} \rangle]. \tag{3.49}$$

Запишем его в виде

$$\delta S_{max} = k_B \sum_k \lambda_k \delta Q_k, \tag{3.50}$$

где «обобщенное тепло» определяется как

$$\delta Q_k = \delta \langle f_i^{(k)} \rangle - \langle \delta f_i^{(k)} \rangle. \tag{3.51}$$

Термин «обобщенное тепло» станет более понятен из дальнейшего, когда мы соприкоснемся с термодинамикой. По аналогии с формулой (3.45) можно указать простое выражение для вариации величины $f_i^{(k)}$:

$$\langle f_i^{(k)2} \rangle - \langle f_i^{(k)} \rangle^2 = \frac{\partial^2 \ln Z}{\partial \lambda_k^2}. \tag{3.52}$$

Во многих практических приложениях $f_i^{(k)}$ зависит еще и от величины α (или от набора таких величин α_1, α_2, ...). Как изменяется среднее значение (3.34) при изменении параметра α? Вычисляя производную от $f_{i,\alpha}^k$ по α и усредняя, получаем

$$\left\langle \frac{\partial f_{i,\alpha}^{(k)}}{\partial \alpha} \right\rangle = \sum_i p_i \frac{\partial f_{i,\alpha}^{(k)}}{\partial \alpha}. \tag{3.53}$$

Подставляя вместо p_i выражение (3.38) и используя равенство (3.41), мы можем записать правую часть соотношения (3.53) в виде

$$\frac{1}{Z} \sum_i \frac{\partial f_{i,\alpha}^{(k)}}{\partial \alpha} \exp\left(-\sum_j \lambda_j f_{i,\alpha}^{(j)}\right). \tag{3.54}$$

В свою очередь это выражение представимо в виде производной от Z по α:

$$(3.54) = -\frac{1}{Z}\frac{1}{\lambda_k}\frac{\partial Z}{\partial \alpha}. \tag{3.55}$$

Так мы приходим к окончательной формуле

$$-\frac{1}{\lambda_k}\frac{\partial \ln Z}{\partial \alpha} = \left\langle \frac{\partial f_{i,\alpha}^{(k)}}{\partial \alpha} \right\rangle. \tag{3.56}$$

Если имеется несколько параметров α_l, то формула (3.56) легко обобщается: необходимо лишь вместо α написать в ней α_l.

Как мы уже неоднократно могли убедиться, величина Z, задаваемая соотношением (3.40), или ее логарифм, очень полезны [см., например, соотношения (3.45), (3.52), (3.56)]. Мы хотим убедиться теперь в том, что соотношение $\ln Z \equiv \lambda$ (см. формулу (3.42)) может быть выведено непосредственно из вариационного принципа. Достаточно взглянуть на соотношение (3.36), чтобы понять возможность следующей его интерпретации. Найдем экстремум выражения

$$\frac{1}{k_B} S - \sum_k \lambda_k \sum_i p_i f_i^{(k)} \tag{3.57}$$

при единственном ограничении

$$\sum_i p_i = 1. \tag{3.58}$$

Из соотношений (3.34), (3.47) и (3.42) следует, что экстремум выражения (3.57) действительно совпадает с $\ln Z$. Заметим, что вариационный принцип для $\ln Z$ по духу отличается от вариационного принципа для S. В случае вариационного принципа для S требовалось найти максимум S при ограничениях (3.34), (3.35), фиксированных f_k и неизвестных λ_k. В случае же вариационного принципа для $\ln Z$ единственным ограничением служит условие (3.58), а величины λ_k предполагаются заданными. Каким образом может быть осуще-

ствлен переход от одного множества заданных величин к другому, станет более понятно после того, как мы рассмотрим в гл. 4 пример из области физики, который заодно позволит пояснить многие другие аспекты излагаемой ниже теории.

3.4. Непрерывные переменные

В большинстве приложений, которые мы имеем в виду, переменная ξ не дискретна, а непрерывна. Но если ξ непрерывна, то, как нетрудно убедиться, информация по крайней мере в общем случае, становится бесконечной. Это связано с тем, что множество состояний образует континуум. Следовательно, нам необходимо кратко обсудить, как можно было бы определить информацию для непрерывных переменных. Начнем с определения плотности вероятности

$$\mathrm{Prob}(\chi \leqslant \xi \leqslant \chi + \Delta\xi) = P(\xi)\Delta\xi. \tag{3.59}$$

Попытаемся теперь учесть, что измерения могут быть произведены только с конечной точностью. Для этого введем погрешность и определим новое распределение вероятности

$$P_\xi(j) = \int_{\xi_j - \varepsilon/2}^{\xi_j + \varepsilon/2} P(\xi)\, d\xi. \tag{3.60}$$

Предполагая, что функция $P(\xi)$ непрерывная, аппроксимируем интеграл (3.60) выражением

$$P_\varepsilon(j) \approx P(\xi_j)\,\varepsilon. \tag{3.61}$$

Информацию при заданной погрешности ε определим как величину

$$I_\varepsilon = -\sum_j P_\varepsilon(j) \ln P_\varepsilon(j). \tag{3.62}$$

Подставляя приближенное выражение (3.61) в формулу (3.62), получаем

$$I_\varepsilon = -\sum_j \varepsilon\, P(\xi_j) \ln P(\xi_j) - \underbrace{\sum_j \varepsilon P(\xi_j)}_{=1} \ln \varepsilon, \tag{3.63}$$

поэтому наш окончательный результат имеет вид

$$I_\varepsilon = -\int d\xi\, P(\xi) \ln P(\xi) - \ln \varepsilon. \tag{3.64}$$

В дальнейшем мы отбросим член $-\ln\varepsilon$, так как он постоянен и не представляет для нас интереса. Обобщение формулы (3.64) на случай нескольких переменных очевидно.

Глава 4
Пример из области физики: термодинамика

Чтобы придать наглядный смысл индексу i, отождествим его со скоростью частицы. В более развитой теории величина p_i есть вероятность заселенности квантового состояния i многочастичной системы. Кроме того, мы отождествим $f_{i,\alpha}^{(k)}$ с энергией E и параметр α с объемом. Таким образом, мы полагаем

$$f_{i,\alpha}^{(k)} = E_i(V); \quad k = 1, \tag{4.1}$$

и устанавливаем взаимно однозначное соответствие между величинами

$$f_1 \leftrightarrow U \equiv \langle E_i \rangle; \quad \alpha \leftrightarrow V; \quad \lambda_1 = \beta. \tag{4.2}$$

В частности, мы положили $\lambda_1 = \beta$. Это позволяет нам записать ряд приведенных ранее формул в таком виде, в котором их можно сразу отождествить с соотношениями, хорошо известными из термодинамики и статистической механики. Например, вместо формулы (3.38) мы получаем соотношение

$$p_i = \exp[-\lambda - \beta E_i(V)], \tag{4.3}$$

задающее знаменитую функцию распределения Больцмана.

Соотношение (3.47) принимает вид

$$\frac{1}{k_B} S_{\max} = \ln Z + \beta U, \tag{4.4}$$

или, после небольших преобразований,

$$U - \frac{1}{k_B} S_{\max} = -\frac{1}{\beta} \ln Z. \tag{4.5}$$

Последнее соотношение хорошо известно в термодинамике и статистической физике. Первый член можно интерпретировать как внутреннюю энергию U, $1/\beta$ – как абсолютную температуру T, умноженную на постоянную Больцмана k_B, $S_{\max}$ – энтропия. Правая часть соотношения (4.5) соответствует свободной энергии $\mathscr{F}$, поэтому в термодинамических обозначениях соотношение (4.5) записывается следующим образом:

$$U - TS = \mathscr{F}. \tag{4.6}$$

Сравнивая обе записи, получаем

$$\mathscr{F} = -k_B T \ln Z \tag{4.7}$$

и $S = S_{max}$. Соответственно мы будем в дальнейшем опускать индекс max. Соотношение (3.40) в новых обозначениях принимает следующий вид:

$$Z = \sum_i e^{-\beta E_i}. \tag{4.8}$$

Это не что иное, как обычная функция распределения. Приведенные нами формулы позволяют без труда вывести ряд других термодинамических соотношений.

Единственное, что требует известной осмотрительности, это деление переменных на независимые и зависимые. Начнем с информационной энтропии S_{max}. В соотношение (3.47) она входит как функция от λ, λ_μ и f_k. Но λ и λ_μ сами определяются уравнениями, содержащими f_k и $f_i^{(k)}$, которые считаются заданными величинами [см. соотношения (3.40), (3.42), (3.43)]. Таким образом, независимыми переменными естественно считать величины f_k и $f_i^{(k)}$, зависимыми – параметры λ и λ_k и тем самым (в силу соотношения (3.47)) S_{max}. На практике $f_i^{(k)}$ – заданные функции от i (например, от энергии i-го состояния), но зависят еще и от параметров α [например, от объема; см. соотношение (4.1)]. Таким образом, при нашем подходе истинно независимыми переменными являются величины f_k (как в приведенном выше случае) и α. В заключение этой главы мы найдем зависимость $S = S(f_k, \alpha)$. В нашем примере $f_1 = E \equiv U$, $\alpha = V$, поэтому

$$S = S(U, V). \tag{4.9}$$

Применим теперь общее соотношение (3.49) к нашей конкретной модели. Если мы варьируем только внутреннюю энергию U, оставляя переменную V фиксированной, то

$$\delta \langle f_i^{(1)} \rangle \equiv \delta f_1 \equiv \delta U \neq 0, \tag{4.10}$$

$$\delta f_{i,\alpha}^{(1)} \equiv \delta E_i(V) = \frac{\delta E_i(U)}{\delta V} \delta U = 0, \tag{4.11}$$

и поэтому

$$\delta S = k_B \lambda_1 \delta U,$$

или

$$\frac{\delta S}{\delta U} = k_B \lambda_1 (\equiv k_B \beta). \tag{4.12}$$

Согласно термодинамике, левая часть соотношения (4.12) определяет обратную абсолютную температуру

$$\frac{\delta S}{\delta U} = \frac{1}{T}. \tag{4.13}$$

Отсюда, как и следовало ожидать, мы получаем, что $\beta = 1/(k_B T)$. С другой стороны, варьируя V, но оставляя фиксированной переменную U. т.е. в случае, если

$$\delta \langle f_i^{(1)} \rangle = 0, \tag{4.14}$$

но

$$\langle \delta f_i^{(1)} \rangle = \langle \frac{\delta E_i(V)}{\delta V} \rangle \delta V \neq 0, \tag{4.15}$$

мы преобразуем соотношение (3.49) к виду

$$\delta S = k_B(-\lambda_1) \langle \frac{\delta E_i(V)}{\delta V} \rangle \delta V,$$

или

$$\frac{\delta S}{\delta V} = -\frac{1}{T} \langle \frac{\delta E_i(V)}{\delta V} \rangle. \tag{4.16}$$

Из термодинамики известно, что

$$\frac{\delta S}{\delta V} = -\frac{P}{T}, \tag{4.17}$$

где P – давление, и, сравнивая формулы (4.17) и (4.16), мы заключаем, что

$$\langle \frac{\delta E_i(V)}{\delta V} \rangle = -P. \tag{4.18}$$

Подставляя выражения (4.13) и (4.17) в соотношение (3.49), приводим его к виду

$$\delta S = \frac{1}{T} \delta U + \frac{1}{T} P \delta V. \tag{4.19}$$

В термодинамике правая часть соотношения (4.19) равна dQ/T, где dQ – тепло. Это объясняет, почему величину (3.51) мы назвали «обобщенным теплом». Приведенные нами соображения могут быть обобщены на случай, когда имеются частицы различного сорта и N_k – заданное среднее число частиц k-го сорта, $k = 1, \ldots, m$. В этом случае мы отождествляем f_1 с E, но $f_{k'+1}$ – с $N_{k'}$; $k' = 1, \ldots, m$ (обратите внимание на сдвиг значения индекса!). Так как каждый (l-й) сорт частиц может быть представлен различными числами N_l, мы производим замену $i \to i, N_1, \ldots, N_m$ и

$$f_i^{(k+1)} \to f_{i, N_1, \ldots, N_m}^{(k+1)} = N_k.$$

Чтобы сохранить соответствие с термодинамикой, мы полагаем

$$\lambda_{k+1} = -\frac{1}{k_B T}\mu_k, \tag{4.20}$$

где величина μ_k называется химическим потенциалом.

Соотношение (3.47) (с учетом соотношения (3.42)) принимает (после умножения его правой и левой части на $k_B T$) следующий вид:

$$TS = k_B T \ln Z + U - \mu_1 \overline{N}_1 - \mu_2 \overline{N}_2 - \ldots - \mu_m \overline{N}_m. \tag{4.21}$$

Соотношение (3.49) позволяет нам записать равенство

$$\frac{\delta S}{\delta \overline{N}_k} = -k_B \lambda_{k+1} = \frac{1}{T}\mu_k. \tag{4.22}$$

Функция распределения имеет вид

$$Z = \sum_{N_1 N_2, \ldots, N_m} \sum_i \exp\left\{-\frac{1}{k_B T}[E_i(V) - \mu_1 N_1 - \ldots - \mu_m N_m]\right\}. \tag{4.23}$$

Приведенные выше соображения применимы главным образом к термодинамике необратимых процессов. Что же касается термодинамики обратимых процессов, то зависимые и независимые переменные здесь могут в известной мере меняться ролями. В нашу задачу не входит обсуждение преобразований переменных, приводящих к различным термодинамическим потенциалам (Гиббса, Гельмгольца и др.). Упомянем лишь один важный случай: вместо U, V (и $N_1, \ldots, N_m$) в качестве независимых переменных можно ввести V, $T = (\partial S/\partial U)^{-1}$ (и $N_1, \ldots, N_m$). Рассмотрим для примера случай, когда в качестве независимых переменных выбраны U и V (при этом мы полагаем формально $\mu_1, \mu_2, \ldots = 0$). Формула (4.7) непосредственно задает свободную энергию $\mathcal{F}$ как функцию от T. Дифференцируя $\mathcal{F}$ по T, получаем

$$-\frac{\partial \mathcal{F}}{\partial T} = k_B \ln Z + \frac{1}{T}\frac{1}{Z}\sum_i E_i e^{-\beta E_i}.$$

Второй член в правой части есть переменная U, поэтому

$$-\frac{\partial \mathcal{F}}{\partial T} = k_B \ln Z + \frac{1}{T}U. \tag{4.24}$$

Сравнивая это соотношение с соотношением (4.5), в котором $1/\beta = k_B T$, приходим к важному результату

$$-\frac{\partial \mathcal{F}}{\partial T} = S, \tag{4.25}$$

где у величины, стоящей в правой части, мы отбросили индекс «max».

Читателям, которые заинтересуются приложением изложенного выше формализма к термодинамике необратимых процессов, т. е. *к релаксационным явлениям*, я рекомендую обратиться к моей книге «Синергетика». В настоящей книге мы сосредоточим усилия совсем в другой области, а именно на неравновесных фазовых переходах в сильно неравновесных физических и биологических системах, а также в нефизических системах (например, в экономике).

Глава 5

Приложение принципа максимума информации к самоорганизующимся системам

5.1. Введение

В гл. 1 мы уже говорили о том, что самоорганизующимися принято называть системы, способные без вмешательства извне образовывать макроскопические пространственные, временные или пространственно-временные структуры. До сих пор функции распределения параметров порядка, управляющих образованием макроскопических структур, можно было вычислять только в рамках микроскопических теорий (см. гл. 2). В этом разделе мы выведем эти функции из макроскопических величин и докажем правильность используемой нами процедуры на примере одно- и многомодового лазера вблизи порога генерации.

Принцип максимума информационной энтропии позволяет делать несмещенные оценки распределения вероятности микроскопических состояний систем, относительно которых в противном случае известны только некоторые средние, соответствующие макроскопическим наблюдениям. Как было показано в предыдущем разделе, принцип максимума информационной энтропии лежит в основе весьма изящного подхода к многим фундаментальным соотношениям и понятиям термодинамики, т. е. вполне примени́м к системам, находящимся в тепловом равновесии. С другой стороны, не были известны сколько-нибудь успешные попытки общего применения принципа максимума информационной энтропии к системам, далеким от теплового равновесия.

В этом разделе я хочу показать, что принцип максимума информационной энтропии может быть вполне успешно применен к неравновесным системам, если в процессе самоорганизации они образуют макроскопические структуры. Попутно мы еще раз получим хорошо известные функции распределения таких систем. Ранее эти функции были получены на основе микроскопических теорий (см. разд. 2.4, 2.7). Однако наш подход позволяет получать и обобщения этих функций. Чтобы пояснить предлагаемую нами процедуру, начнем с примера – одномодового лазера.

5.2. Приложение к самоорганизующимся системам: одномодовый лазер

Самоорганизующиеся системы характеризуются способностью образовывать макроскопические структуры, описываемые адекватными

параметрами порядка. Мы предпочитаем вместо абстрактных общих соображений проиллюстрировать предлагаемую нами процедуру на нескольких конкретных примерах. Очень важно строго придерживаться основной идеи принципа максимума информационной энтропии, а именно рассматривать макроскопически наблюдаемые величины.

В случае одномодового лазера экспериментально наблюдаемой величиной являются интенсивность излучения и второй момент интенсивности в случае стационарного состояния. Из теории лазера хорошо известно, что основное различие между излучением лампы и лазера становится очевидным только после того, как помимо первых моментов интенсивности мы измерим вторые моменты. Измерению подлежат также корреляции интенсивности, но поскольку мы имеем в виду стационарную теорию, информацией, содержащейся в корреляциях, можно пренебречь.

Зависящую от пространственных координат и времени напряженность электрического поля одномодового лазера можно записать в виде

$$E(x,\ t) = E(t)\sin kx, \tag{5.1}$$

где амплитуда $E(t)$ может быть разложена на части, соответствующие положительным и отрицательным частотам:

$$E(t) = Be^{-i\omega t} + B^* e^{i\omega t}. \tag{5.2}$$

Если измерить интенсивность лазерного излучения на интервалах времени, больших по сравнению с периодом колебаний, но малых по сравнению с временами флуктуаций $B(t)$, то интенсивность на выходе окажется пропорциональной B^*B и скорости потерь $2\varkappa$ лазера. Для простоты мы отбросим все остальные постоянные и запишем

$$I = 2\varkappa B^* B. \tag{5.3}$$

Аналогичным образом квадрат интенсивности при усреднении по тому же временному интервалу оказывается равным

$$I^2 = 4\varkappa^2 B^{*2} B^2. \tag{5.4}$$

Вследствие флуктуаций лазера B и B^* – случайные переменные, соответствующие стационарному процессу. Это позволяет нам отождествить B и B^* с индексами i величин p_i в соотношении (3.32), где мы положили $k_B = 1$ и интерпретировали правую часть как информацию i. Поскольку случайные переменные B теперь не дискретные, а непрерывные, мы должны заменить суммирование по индексам i интегрированием:

$$i = -\int p(B,B^*)\ln p(B,B^*)\,d^2B. \tag{5.5}$$

Соотношение (3.34) может быть интерпретировано как интегралы по d^2B с вероятностью $p(B,B^*)$ в качестве весовой функции. Обозначая

средние угловыми скобками, приходим к следующим двум ограничениям:

$$f_1 = \langle 2\varkappa B^* B \rangle, \tag{5.6}$$

$$f_2 = \langle 4\varkappa^2 B^{*2} B^2 \rangle, \tag{5.7}$$

Кроме того, руководствуясь той же аналогией, мы приходим к необходимости определить $f_i^{(k)}$ соотношениями

$$f^{(1)}_{B.B^*} = 2\varkappa B^* B, \tag{5.8}$$

$$f^{(2)}_{B.B^*} = 4\varkappa^2 B^{*2} B^2. \tag{5.9}$$

Теперь мы уже в состоянии применить непосредственно формулу (3.38) и получить

$$p(B,B^*) = \exp[-\lambda - \lambda_1 2\varkappa B^* B - \lambda_2 4\varkappa^2 (B^* B)^2]. \tag{5.10}$$

В несколько иных обозначениях соотношение (5.10) имеет вид

$$p(B,B^*) = N \exp(-\alpha |B|^2 - \beta |B|^4). \tag{5.11}$$

Эта функция хорошо известна в лазерной физике. Она была выведена Рискеном путем решения уравнения Фоккера–Планка, соответствующего уравнению Ланжевена для лазера, выведенному ранее Г. Хакеном.

Заметим, что в случае лазера коэффициент α должен быть отрицательным. Но вблизи порога α может принимать и отрицательные, и положительные значения.

5.3. Многомодовый лазер без фазовых соотношений

В этом случае напряженность поля разлагается на моды по формуле

$$E(x,\ t) = \sum_l E(t) \sin k_l x, \tag{5.12}$$

где для простоты мы рассматриваем только аксиальные моды. Амплитуды мод так же, как и в предыдущей задаче, могут быть разложены на части, соответствующие положительным и отрицательным частотам:

$$E_l(t) = B_l(t) e^{-i\omega_l t} + B_l^*(t) e^{i\omega_l t}. \tag{5.13}$$

Интенсивность, усредненная по интервалам времени, большим по сравнению с периодом колебаний и малым по сравнению с периодами флуктуаций, определяется выражением

$$I_l \sim 2\varkappa_l B_l^* B_l \sim n_l. \tag{5.14}$$

Результаты, полученные в разд. 5.2, допускают непосредственное обобщение, если n_l, или, что эквивалентно, B_l^* и B_l мы будем считать стохастическими переменными. Получаем

$$f_l = \langle n_l \rangle, \tag{5.15}$$

$$f_{l,l'} = \langle n_l, n_{l'} \rangle \tag{5.16}$$

и, отождествляя k и i и устанавливая соответствия

$$k \leftrightarrow \begin{cases} l \\ l, l' \end{cases};$$

$$i \leftrightarrow (n_1, n_2, \ldots, n_M) = \mathbf{n}, \tag{5.17}$$

приходим к соотношениям

$$f_{\mathbf{n}}^{(l)} = n_l, \tag{5.18}$$

$$f_{\mathbf{n}}^{(l,l')} = n_l n_{l'}, \tag{5.19}$$

$$p_i \to p(\mathbf{n}). \tag{5.20}$$

Применяя формулу (3.38), получаем

$$p(\mathbf{n}) \equiv p(n_1, \ldots, n_M) = \exp\left(-\lambda - \sum_l \lambda_l n_l - \sum_{l\,l'} \lambda_{l,l'} n_l n_{l'}\right). \tag{5.21}$$

Это конечный результат наших вычислений. Правую часть соотношения (5.21) можно представить в эквивалентном виде

$$N \exp\left(-\sum_l \alpha_l n_l - \sum_{ll'} \beta_{ll'} n_l n_{l'}\right). \tag{5.22}$$

Именно в такой форме она была получена из многомодового уравнения Фоккера–Планка для частных случаев, когда решение может быть в явном виде построено с помощью принципа детального равновесия.

5.4. Процессы, периодические по параметрам порядка

Мы хотим показать, каким образом можно найти функцию распределения, если рассматриваемые процессы периодичны по параметрам порядка.

Рассмотрим в качестве конкретного примера угловую координату ϕ, и пусть моменты периодичны по ϕ с периодом 2π. Вместо моментов ϕ в этом случае разумно рассматривать соответствующие моменты периодических функций, т. е.

$$\langle \sin\phi \rangle, \langle \cos\phi \rangle, \ldots, \tag{5.23}$$

или, более общо,

$$\langle \sin n\phi \rangle, \langle \cos n\phi \rangle, \tag{5.24}$$

где n – целое положительное число.

Чтобы продемонстрировать нашу процедуру, рассмотрим частный случай, в котором по соображениям симметрии

$$\langle \sin n\phi \rangle = 0 \text{ при всех } n. \tag{5.25}$$

При нашем подходе, как упоминалось выше, мы сохраняем только несколько первых моментов. Если по аналогии с этим мы сохраним только два члена в системе функций (5.24), то получим

$$p(\phi) = \exp(\lambda + \lambda_1 \cos\phi + \lambda_2 \cos 2\phi) = \exp[V(\phi)]. \tag{5.26}$$

Как было нами показано на конкретных примерах, принцип максимума информационной энтропии дает прямой способ получения в общем виде функций распределения для неравновесной системы (такой, как лазер). Полученные результаты согласуются с функциями распределения, вычисленными в рамках микроскопических теорий при некоторых ограничительных условиях. Теперь совершенно ясно, как следует поступать в других случаях. Полное состояние системы $\mathbf{q}(\mathbf{x}, t)$ следует спроектировать на функции, описывающие наблюдаемую макроскопическую пространственную или временную структуру. Так мы найдем амплитуды, моменты которых доступны измерению и тем самым могут служить функциями, определяющими f_k и $f_i^{(k)}$.

Несмотря на успешное приложение принципа максимума информационной энтропии, не следует забывать о том, что о множителях Лагранжа нам известно мало и их, разумеется, следует определять экспериментально. С другой стороны, преимущество микроскопической теории заключалось в том, что эти константы определялись из первых принципов, и поэтому, в частности, можно было предсказать, что параметр α_l меняет знак в точках неустойчивости. Вполне возможно, однако, что такие же свойства удастся вывести и в рамках макроскопической теории.

Наш подход вполне применим к целому ряду физических проблем: конвекции в жидкостях, образованию структур в химических реакциях и росту в морфогенетических полях.

Между нашими результатами и результатами термодинамики обратимых процессов существует формальное сходство. Однако нельзя не отметить и существенные различия между теми и другими результатами. Прежде всего следует иметь в виду, что физический смысл констант в нашей теории сильно отличается от физического смысла констант в системах, находящихся в тепловом равновесии. Например, в таких неравновесных системах, как лазеры, нам приходится иметь дело с выходными интенсивностями, тогда как в случае равновесных систем аналогичную роль играют, например, энергии. Это также отчетливо проявляется и в рассмотрении на микроскопическом уровне: если в микроскопической теории равновесных систем решающую роль играют энергии, то в неравновесных систе-

мах образование структур определяется константами скоростей и скоростями роста.

Мы можем сделать несколько далеко идущих выводов. До сих пор принцип максимума информации (информационной энтропии) применялся в термодинамике обратимых и необратимых процессов, но не в теории неравновесных фазовых переходов. В термодинамике как обратимых, так и необратимых процессов в число ограничений входит ограничение по энергии. В нашем случае мы имеем ограничение по *выходной интенсивности* I. Еще важнее то, что нам необходимо включать в теорию корреляцию интенсивности в форме второго момента, т. е. $\langle I^2 \rangle$. В равновесной термодинамике, когда речь идет о принципе максимума энтропии, так никогда не поступают. Но, как мы теперь знаем, включение второго момента $\langle I^2 \rangle$ совершенно естественно и необходимо для неравновесных фазовых переходов. При достижении порога или выше его выполняется неравенство $\lambda_1 \geqslant 0$, и интеграл от $\exp(\lambda_1 |b|^2)$ расходится, что соответствует критическим флуктуациям. По этой причине и чтобы учесть их ограниченность вследствие насыщения, как раз и приходится учитывать $\langle I^2 \rangle$. Иначе говоря, вблизи неравновесных фазовых переходов флуктуации становятся «наблюдаемыми», и их следует учитывать путем введения адекватных ограничений. Для меня представляется несомненным, что то же верно и для фазовых переходов в системах, находящихся в тепловом равновесии, и что в этом случае требуется обобщение принципа максимума информационной энтропии.

Глава 7
Информация, приращение информации и эффективность самоорганизующихся систем вблизи точек неустойчивости

7.1. Введение

В этом разделе мы хотим продолжить исследование результатов, полученных в предыдущей главе. Мы показали, что функция распределения совместной вероятности $P(\xi_u, \xi_s)$, относящаяся к параметрам порядка ξ_u и амплитудам ξ_s подчиненных мод, допускает разложение в произведение вида

$$P(\xi_u, \xi_s) = \prod_s P_s(\xi_s \mid \xi_u) f(\xi_u). \tag{7.1}$$

Это разложение соответствует принципу подчинения микроскопической теории. Теперь мы хотим показать, что разложение (7.1) позволяет разложить информацию и приращение информации на части, одна из которых соответствует только параметрам порядка, а другая представляет собой сумму информаций подчиненных мод, усредненных по распределению параметров порядка. Как мы увидим, вблизи точек неустойчивости информация, относящаяся к параметрам порядка, резко изменяется, в то время как информация, относящаяся к подчиненным модам, резкого изменения не испытывает. Следовательно, вблизи точек неустойчивости достаточно исследовать поведение информации, относящейся к параметрам порядка, и приращения информации, что и проделано нами в этой главе для широкого класса систем, претерпевающих неравновесные фазовые переходы. Кроме того, будет показано, что информация, приращение информации и эффективность (определяемая в смысле, указываемом в этом разделе) допускают непосредственное измерение.

Информацией по определению называется величина

$$i = -\sum_j p_j \ln p_j, \tag{7.2}$$

а приращением информации – величина

$$K = \sum_j p_j \ln \frac{p_j}{p'_j}. \tag{7.3}$$

Здесь p_j – распределение вероятности состояний, характеризуемых индексом j; p_j и p'_j в формуле (7.3) – два различных распределения вероятности. Чтобы установить более тесный контакт с полученными ранее результатами, мы отождествим индекс j с значениями, принимаемыми стохастическими переменными ξ_u, ξ_s. Как обычно, мы

будем понимать под ξ_u, ξ_s не только сами переменные, но и их допустимые значения. Заметим, что в силу соотношения (2.78) вектор состояния $\mathbf{q}$ определен, если известны ξ_u и ξ_s.

7.2. Принцип подчинения и его применение к информации

Мы хотим показать, что разложение (7.1) позволяет особым образом записать информацию (7.2). Для этого подставим разложение (7.1) в соотношение (7.2). Заменяя логарифм произведения суммой логарифмов, получаем

$$i = -\sum_{\xi_u, \xi_s} \prod_s P_s(\xi_s \mid \xi_u) f(\xi_u) \left[\ln f(\xi_u) + \sum_s \ln P_s(\xi_s \mid \xi_u)\right]. \tag{7.4}$$

Используя условие нормировки

$$\sum_{\xi_s} P_s(\xi_s \mid \xi_u) = 1, \tag{7.5}$$

мы можем записать выражение (7.4) в виде

$$i = -\sum_{\xi_u} f(\xi_u) \ln f(\xi_u) - \sum_{\xi_u, \xi_s, s} f(\xi_u) P_s(\xi_s \mid \xi_u) \ln P_s(\xi_s \mid \xi_u). \tag{7.6}$$

Последнее соотношение удобно представить следующим образом:

$$i = I_f + \sum_{s\xi_u} f(\xi_u) I_s(\xi_u), \tag{7.7}$$

где

$$I_f = -\sum_{\xi_u} f(\xi_u) \ln f(\xi_u) \tag{7.8}$$

– информация, относящаяся к параметрам порядка, в то время как

$$I_s = -\sum_{\xi^s} P_s(\xi_s \mid \xi_u) \ln P_s(\xi_s \mid \xi_u) \tag{7.9}$$

– информация, относящаяся к подчиненной подсистеме или подчиненной моде с индексом s. Как будет показано ниже, вблизи точек неустойчивости информация I_f изменяется сильно, а информация I_s – слабо.

Как нетрудно видеть, величина I_s является информацией в предположении, что величина ξ_u принимает определенное значение. В интересующем нас плане это означает, что параметр порядка подчиняет моды особым образом, который гарантирует возникновение макроскопической структуры через самоорганизацию.

7.3. Приращение информации

Распределения, соответствующие величинам p и p', условимся различать по индексам n и a. Тогда по определению (7.3) приращение информации определяется выражением

$$K = \sum_{\xi_u,\xi_s} P_n(\xi_u | \xi_s) \ln \frac{P_n(\xi_u,\xi_s)}{P_a(\xi_u,\xi_s)}. \tag{7.10}$$

Используя разложение (7.1), запишем соотношение (7.10) в виде

$$K = \sum_{\xi_u,\xi_s} f_n \prod_s P_{s,n} \left[\ln f_n + \sum_{s'} \ln P_{s',n} - \ln f_a - \sum_{s'} \ln P_{s',a} \right], \tag{7.11}$$

где мы снова заменили логарифм произведения суммой логарифмов. Условие нормировки (7.5) позволяет после несложной перегруппировки членов представить K в виде суммы

$$K = \sum_{\xi_u} f_n \ln \frac{f_n}{f_a} + \sum_s \sum_{\xi_u} f_n \sum_{\xi_s} P_{s,n} \ln \frac{P_{s,n}}{P_{s,a}}, \tag{7.12}$$

которой можно придать окончательную форму

$$K = K_f + \sum_s \sum_{\xi_u} f_n K_s, \tag{7.13}$$

где K_f – величина, определяемая выражением

$$K_f = \sum_{\xi_u} \ln \frac{f_n}{f_a}, \tag{7.14}$$

а K_s – величина, определяемая выражением

$$K_s = \sum_{\xi_s} P_{s,n}(\xi_s | \xi_u) \ln \frac{P_{s,n}(\xi_s | \xi_u)}{P_{s,a}(\xi_s | \xi_u)}. \tag{7.15}$$

7.4. Пример: неравновесные фазовые переходы

В этом разделе мы хотим продемонстрировать полезность выведенных нами в предыдущих разделах формул на примере неравновесных фазовых переходов (см. разд. 2.7). Будем предполагать, что собственное число λ_u вещественно. Для большей наглядности выпишем типичные эволюционные уравнения для системы, претерпевающей неравновесные фазовые переходы. В обозначениях параметров порядка ξ_u и амплитуд подчиненных мод ξ_s эти уравнения представимы в виде

$$\dot{\xi}_u = \lambda_u \xi_u + h_u(\xi_u, \xi_s) + F_u(t), \tag{7.16}$$

$$\dot{\xi}_s = \lambda_s \xi_s + g_s(\xi_u) + \xi_s k_s(\xi_u) + \ldots + F_s(t). \tag{7.17}$$

Если система управляется извне некоторым (управляющим) параметром, то все величины, стоящие в правых частях уравнений (7.16) и (7.17), зависят от этого управляющего параметра, но зависят по-разному.

Когда мы нормируем управляющий параметр так, чтобы неустойчивости соответствовало значение параметра $\alpha = 0$, то λ_u и λ_s зависят от α следующим образом:

$$\lambda_u \sim \alpha^{\varkappa}, \lambda_s = \lambda_s(0) + O(\alpha) \approx \lambda_s(0), \tag{7.18}$$

где $\varkappa$ – некоторое положительное число. Ясно, что λ_u очень чувствительно к малейшим изменениям параметра α, в то время как λ_s зависит от α лишь очень слабо, так как главный член разложения является отличной от нуля константой. Аналогичным образом функции h_u, g_s, k_s, F_u и F_s лишь слабо зависят от α. Чтобы придать нашему примеру как можно большую наглядность, применим принцип подчинения в его главной аппроксимации, порождающей члены, которые получаются методом адиабатического исключения переменных, использующим приближенное равенство

$$\dot{\xi}_s = 0. \tag{7.19}$$

Это позволяет найти главные члены решения уравнения (7.17)

$$\xi_s \approx -\frac{1}{\lambda_s(0)} g_s(\xi_u) - \frac{1}{\lambda_s(0)} F_s(t). \tag{7.20}$$

В разд. 2.6 (см. соотношение (2.91)) мы вывели соответствующее распределение условной вероятности $P_s(\xi_s | \xi_u)$. Оно имеет следующий вид:

$$P_s(\xi_s | \xi_u) = N \exp\left\{ -\left[\xi_s + \frac{g_s(\xi_u)}{\lambda_s} \right]^2 Q^{-1} \right\}. \tag{7.21}$$

Теперь мы можем подставить это выражение в формулу (7.9), задающую I_s. Но, вводя новую переменную

$$\xi_s' = \xi_s - \frac{g_s(\xi_u)}{|\lambda_s|}, \tag{7.22}$$

мы получаем возможность исключить зависимость распределения вероятности подчиненных переменных от ξ_u, так что информация I_s перестает зависеть от ξ_u. Следовательно, в выражении (7.7) мы можем выполнить интегрирование по ξ_u во втором члене. В результате мы получаем

$$I = I_f + \sum_s I_s, \tag{7.23}$$

где вторая часть (по крайней мере в рассматриваемом приближении) не зависит от α. Таким образом, вблизи точки неустойчивости изменение информации зависит только от изменения параметра порядка:

$$I(\alpha_1) - I(\alpha_2) \approx I_f(\alpha_1) - I_f(\alpha_2). \tag{7.24}$$

Используя приближение того же типа, мы можем записать приращение информации в виде

$$K = K_f = \int d^n \xi_u f(\xi_u, \alpha_2) \ln \frac{f(\xi_u, \alpha_2)}{f(\xi_u, \alpha_1)}. \tag{7.25}$$

В дальнейшем мы подробно исследуем I_f. Затем в разд. 7.11 мы изучим зависимость I_s от α. Как будет показано, I_s монотонно изменяется вместе с параметром λ, тогда как в зависимости I_f от α обнаруживается своего рода «сингулярность».

7.5. Мягкие одномодовые неустойчивости

Как было показано в предыдущем разделе, изменения информации и приращения информации всей системы вблизи точек неустойчивости практически совпадают соответственно с изменениями информации и приращения информации, относящимися к параметрам порядка. В этом разделе мы хотим вычислить i и K, используя знание функции распределения, относящейся к параметру порядка (по образцу одного параметра порядка, претерпевающего неравновесный фазовый переход второго порядка). Как было показано в разд. 2.7, функция распределения f имеет следующий вид:

$$f(\xi) = N \exp(\alpha\xi^2 - \beta\xi^4). \tag{7.26}$$

Подставляя экспоненту (7.26) в определение информации (7.8), получаем

$$I_f = -\ln N - \alpha\langle\xi^2\rangle + \beta\langle\xi^4\rangle, \tag{7.27}$$

где для краткости мы воспользовались обозначением

$$\langle\xi^n\rangle = \int_{-\infty}^{+\infty} f(\xi)\xi^n d\xi. \tag{7.28}$$

Аналогичным образом приращение информации (7.14) принимает вид

$$K = (f_\alpha, f_{\alpha_0}) = \int f_\alpha \ln \frac{f_\alpha}{f_{\alpha_0}} d\xi = \ln N(\alpha) - \ln N(\alpha_0) + (\alpha - \alpha_0)\langle\xi^2\rangle_\alpha. \tag{7.29}$$

7.6. Как можно измерить информацию и приращение информации?

7.6.1. Эффективность

Изучая поведение самоорганизующихся систем, например лазера, исследователи установили, что мерой макроскопического действия таких систем может служить квадрат параметра порядка.

Например, в случае лазера в качестве параметра порядка выступает мода поля, и действие лазера может быть измерено квадратом амплитуды моды поля. Это наводит на мысль о необходимости введения величины

$$\Omega(\alpha) = \langle \xi_u^2 \rangle \tag{7.30}$$

в качестве меры макроскопического действия, или, точнее, работы, производимой системой. Среднее определяется величиной

$$\langle \xi^2 \rangle = \int \xi^2 f(\xi)\, d\xi, \tag{7.31}$$

где мы используем явный вид функции распределения

$$f(\xi) = N \exp(\alpha\xi^2 - \beta\xi^4) \tag{7.32}$$

с нормировочным множителем, определяемым выражением

$$N^{-1} = \int_{-\infty}^{+\infty} \exp(\alpha\xi^2 - \beta\xi^4)\, d\xi. \tag{7.33}$$

Определим теперь эффективность W как скорость изменения величины Ω при изменении управляющего параметра α, поскольку в ряде самоорганизующихся систем параметр α связан с входной мощностью. Именно поэтому мы будем пока интерпретировать параметр α просто как входную мощность:

$$W = \frac{d\Omega}{d\alpha}. \tag{7.34}$$

Из соотношений (7.31)–(7.33) следует, что

$$W = \frac{\int \xi^4 \exp(\alpha\xi^2 - \beta\xi^4)\, d\xi}{\int \exp(\alpha\xi^2 - \beta\xi^4)\, d\xi} - \frac{[\int \exp(\alpha\xi^2 - \beta\xi^4)\xi^2\, d\xi]^2}{[\int \exp(\alpha\xi^2 - \beta\xi^4)\, d\xi]^2}. \tag{7.35}$$

Последнее соотношение представимо в виде

$$W = \langle \xi_u^4 \rangle - \langle \xi_u^2 \rangle^2 \tag{7.36}$$

(мы записали ξ_u вместо ξ). Если управляющий параметр α входит в функцию распределения (7.32) неявно:

$$\exp[\lambda(\alpha)\xi_u^2 - \beta\xi_u^4], \tag{7.37}$$

то соотношение (7.36) необходимо заменить следующим:

$$W = \frac{d\lambda}{d\alpha}(\langle \xi_u^4 \rangle - \langle \xi_u^2 \rangle^2). \tag{7.38}$$

Вычисление величины W выше и ниже точки неустойчивости не составляет труда, и в дальнейшем мы приведем соответствующие результаты.

7.6.2. Информация и приращение информации

Займемся теперь вычислением информации I_f. Из формулы (7.32) с учетом соотношения (7.33) получаем

$$-I_f = \ln N + \alpha\langle\xi^2\rangle - \beta\langle\xi^4\rangle. \quad (7.39)$$

Хотя второй и четвертый моменты могут быть измерены непосредственно, тем не менее $\ln N$ следует связать с измеримыми величинами.

Для этого мы прежде всего продифференцируем $\ln N$ по α:

$$\frac{d\ln N}{d\alpha} = -\frac{\int \exp(\alpha\xi^2 - \beta\xi^4)\xi^2 d\xi}{\int \exp(\alpha\xi^2 - \beta\xi^4) d\xi}. \quad (7.40)$$

Нетрудно видеть, что это соотношение приводится к виду

$$\frac{d\ln N}{d\alpha} = -\langle\xi^2\rangle. \quad (7.41)$$

Интегрируя уравнение (7.41), получаем

$$\ln N = -\int_{\alpha_0}^{\alpha} \langle\xi^2\rangle d\alpha + \ln N(\alpha_0). \quad (7.42)$$

Подставляя это соотношение в выражение (7.39), преобразуем последнее к виду

$$-I_f = -\int_{\alpha_0}^{\alpha} \langle\xi^2\rangle d\alpha + \ln N(\alpha_0) + \alpha\langle\xi^2\rangle - \beta\langle\xi^4\rangle. \quad (7.43)$$

Чтобы исключить $\ln N$, вычислим соответствующую информацию при α_0 и получим окончательный результат

$$-I_f(\alpha) + I_f(\alpha_0) = -\int_{\alpha_0}^{\alpha} \langle\xi^2\rangle d\alpha + \alpha\langle\xi^2\rangle_\alpha - \beta\langle\xi^4\rangle_\alpha - \alpha_0\langle\xi^2\rangle_{\alpha_0} + \beta\langle\xi^4\rangle_{\alpha_0}, \quad (7.44)$$

или в дифференциальной форме

$$-\frac{dI_f}{d\alpha} = -\alpha\langle\xi^2\rangle^2 + \alpha\langle\xi^4\rangle + \beta\langle\xi^2\rangle\langle\xi^4\rangle - \beta\langle\xi^6\rangle. \quad (7.45)$$

Аналогично приращение информации может быть записано следующим образом:

$$K = -\int_{\alpha_0}^{\alpha} \langle\xi^2\rangle_\alpha d\alpha' + (\alpha - \alpha_0)\langle\xi^2\rangle_\alpha. \quad (7.46)$$

7.7. Случай нескольких параметров порядка

Теперь мы хотим обобщить результаты, полученные в разд. 7.6, на случай нескольких параметров порядка. Предположим, что функция распределения f имеет вид

$$f(\xi_u) = N \exp\left[\sum_k \Delta_k(\alpha)\xi_k^2 + h(\xi_u)\right]. \quad (7.47)$$

В ряде случаев именно такую функцию распределения удается вывести исходя из микроскопических моделей (см. разд. 2.4) или из принципа максимума информации (см. гл. 6).

В этом разделе мы изложим результаты относительно приращения информации, которые окажутся нечувствительными к конкретной форме нелинейных функций h, если те не зависят от управляющих параметров. В действительности функции h все же зависят от параметра α, но, как правило, слабо. Подставляя выражение (7.47) в определение величины K_f, получаем

$$K_f = \int f \ln \frac{f}{f_0} d^n\xi = \int N \exp\left[\sum_k \Delta_k(\alpha)\xi_k^2 + h(\xi_u)\right] \times \\ \times \left\{\ln N - \ln N_0 + \sum_k [\Delta_k(\alpha) - \Delta_k(\alpha_0)]\xi_k^2\right\} d^n\xi. \tag{7.48}$$

Это выражение представимо в виде

$$K = \ln N - \ln N_0 + \sum_k [\Delta_k(\alpha) - \Delta_k(\alpha_0)] \langle \xi_k^2 \rangle_\alpha. \tag{7.49}$$

Остается вычислить $\ln N$. Продифференцируем для этого $\ln N$:

$$\frac{d \ln N}{d\alpha} = -\frac{\sum_k (\partial\Delta_k/\partial\alpha) \int \xi_k^2 \, d^n\xi \exp(\ldots)}{\int d^n \xi \exp(\ldots)}. \tag{7.50}$$

Полученное соотношение представимо в виде

$$\frac{d \ln N}{d\alpha} = -\sum_k \frac{\partial \Delta_k}{\partial \alpha} \langle \xi_k^2 \rangle. \tag{7.51}$$

Интегрируя уравнение (7.51) по α и подставляя результат интегрирования в соотношение (7.49), приходим к окончательному выражению:

$$K = -\int_{\alpha_0}^{\alpha} \sum_k \frac{\partial \Delta_k}{\partial \lambda} \langle \xi_k^2 \rangle_\alpha d\alpha + \sum_k [\Delta_k(\alpha) - \Delta_k(\alpha_0)] \langle \xi_k^2 \rangle_\alpha. \tag{7.52}$$

В заключение мы вычислим эффективность, связанную с параметрами порядка ξ_l. Ее выходное значение определяется величиной

$$\Omega(\alpha) = \langle \xi_l^2 \rangle. \tag{7.53}$$

Как показывают несложные вычисления, эффективность определяется выражением

$$W = \frac{d\Omega}{d\alpha} = \sum_k \frac{\partial \Delta_k}{\partial \alpha} (\langle \xi_l^2 \xi_k^2 \rangle - \langle \xi_l^2 \rangle \langle \xi_k^2 \rangle). \tag{7.54}$$

В этом разделе мы показали, что изменение информации, приращение информации и эффективность – величины, которые могут быть выражены через параметры порядка. Это проливает новый свет

на поведение самоорганизующихся систем, которые относительно только что упомянутых величин ведут себя так, как если бы они управлялись небольшим числом степеней свободы. В частности, было показано, что информация, приращение информации и эффективность – измеримые величины. Можно было бы ожидать, что они измеримы так же, как измерима энтропия системы, находящейся в тепловом равновесии. Но внутренний механизм, порождающий порядок в неравновесных системах, сильно отличается от механизма, устанавливающего порядок в системах, находящихся в тепловом равновесии. Различен и физический смысл термодинамической энтропии и информации в этих двух случаях.

7.8. Вычисление информации в случае одного параметра порядка

Вычислим теперь информацию для одного параметра порядка вблизи неравновесного фазового перехода, т. е. величину (7.27). Затем, опираясь на полученный результат, мы обсудим, почему соответствующие величины интерпретируются как информация, а не как энтропия.

Напомним, что параметр порядка непрерывен. В этом случае информацию следует определять как

$$I_f = -\int d\xi f(\xi) \ln f(\xi) - \ln \varepsilon, \tag{7.55}$$

где ε – величина интервала (в безразмерных единицах), задающего погрешность измерений.

Зная функцию распределения (7.26), попытаемся вычислить величину (7.27), т. е.

$$I_f = -\ln N - \alpha \langle \xi^2 \rangle + \beta \langle \xi^4 \rangle - \ln \varepsilon. \tag{7.56}$$

Ниже и выше порога величина (7.56) вычисляется просто.

7.8.1. Ниже порога

В этой области параметр α отрицателен, и его абсолютная величина очень велика, поэтому в соотношении (7.26) существен только член, квадратичный по ξ. Соответственно нам необходимо вычислить информацию (7.56) для гауссовской функции

$$f \approx N \exp(-|\alpha| \xi^2). \tag{7.57}$$

Нормировочный множитель мы находим без труда: он оказывается равным

$$N = \sqrt{\frac{|\alpha|}{\pi}}. \tag{7.58}$$

Входящие в выражение (7.56) интегралы легко вычисляются, и мы, в частности, получаем

$$|\alpha|\langle\xi^2\rangle = \frac{1}{2}, \tag{7.59}$$

$$\beta|\xi^4| = \frac{3\beta}{(2|\alpha|)^2}. \tag{7.60}$$

При достаточно больших $|\alpha|$ величиной (7.60) по сравнению с величиной (7.59) можно пренебречь. Окончательный результат оказывается следующим:

$$I_f = -\frac{1}{2}\ln|\alpha| + \frac{1}{2} + \frac{1}{2}\ln\pi - \ln\varepsilon. \tag{7.61}$$

Рассмотрим теперь, как ведут себя $\ln|\alpha|$ и $\ln\varepsilon$. Так как ε – погрешность измерения, можно предположить, что ε – величина достаточно малая и что

$$\varepsilon \to 0, \quad -\ln\varepsilon \to +\infty. \tag{7.62}$$

С другой стороны, так как система находится существенно ниже порога, можно предположить, что

$$|\alpha| \to \infty, \quad -\frac{1}{2}\ln|\alpha| \to -\infty. \tag{7.63}$$

Если фиксировать ε и увеличивать $|\alpha|$, то информация I_f сначала обращается в нуль, а затем становится отрицательной, что невозможно по математическим причинам, так как информация – величина неотрицательная. Но гораздо важнее другое: существует физическая причина, по которой величина $|\alpha|$ не может превосходить некоторое значение. Дело в том, что функция распределения (7.26) существенно отличается от нуля только на интервале ε, задающем погрешность измерения. Это означает, что различение состояний в пределах интервала ε утрачивает смысл. То же можно сформулировать и несколько иначе: мы можем почти с достоверностью утверждать, что система находится в некотором состоянии внутри интервала ε. Остается неоднозначность относительно выбора фиксированного значения параметра α, т. е. относительно того, что означает «почти с достоверностью». Например, мы можем потребовать, чтобы выполнялось равенство

$$I_f = 0, \tag{7.64}$$

из которого следует условие на α:

$$-\frac{1}{2}\ln|\alpha| + \frac{1}{2} + \frac{1}{2}\ln\pi - \ln\varepsilon = 0. \tag{7.65}$$

Так как нас интересуют только изменения информации, мы можем выбирать нуль информации произвольно, поэтому ничто не мешает нам, в частности, потребовать, чтобы

$$I_f = \frac{1}{2} + \frac{1}{2}\ln\pi. \tag{7.66}$$

При таком выборе нуля соотношение между α и ε становится особенно простым, а именно:

$$-\frac{1}{2}\ln|\alpha| - \ln\varepsilon = 0. \tag{7.67}$$

Из него мы с легкостью получаем

$$|\alpha|^{1/2} = \frac{1}{\varepsilon}. \tag{7.68}$$

7.8.2. Выше порога

Как показывает даже поверхностное исследование функции распределения (7.26), ее при достаточно больших α можно аппроксимировать гауссовскими функциями. Для этого мы воспользуемся методом перевала и запишем выражение (7.26) в виде

$$f(\xi) = N\exp[g(\xi)], \tag{7.69}$$

где, разумеется,

$$g(\xi) = \alpha\xi^2 - \beta\xi^4. \tag{7.70}$$

Экстремум функции (7.70) и, следовательно, функции f определяется условием

$$g'(\xi) = 0 = 2\alpha\xi - 4\beta\xi^3. \tag{7.71}$$

Решая уравнение (7.71), находим его ненулевые решения

$$\xi_0 = \pm\sqrt{\frac{\alpha}{2\beta}}. \tag{7.72}$$

Выберем пока знак «плюс». Введем новую переменную η, определяемую соотношением

$$\xi = \xi_0 + \eta. \tag{7.73}$$

Разложение

$$g(\xi) = g(\xi_0 + \eta) = g(\xi_0) + g'(\xi_0)\eta + \frac{1}{2}g''(\xi_0)\eta^2 \tag{7.74}$$

позволяет записать функцию (7.69) в виде

$$f = N\exp\left(\frac{\alpha^2}{4\beta^2} - 2\alpha\eta^2\right), \tag{7.75}$$

где нормировку следует выбирать в интервале $-\infty < \xi < +\infty$. Тогда нормировочный множитель определяется выражением

$$N = \frac{2\alpha}{\pi}\exp\left(-\frac{\alpha^2}{4\beta}\right). \tag{7.76}$$

До сих пор мы рассматривали случай, когда функция f имеет только один максимум. Оценим сначала информацию для случая, когда может реализоваться только один максимум, т.е. когда мы искусственно нарушаем симметрию. Тогда информация I_f легко вычисляется и оказывается, что

$$I_f = \frac{1}{2}\ln\alpha - \frac{1}{2}\ln 2 + \frac{1}{2}\ln\pi + \frac{1}{2} - \ln\varepsilon. \tag{7.77}$$

Обратимся теперь к случаю, представляющему для нас основной интерес, когда полная функция распределения имеет два максимума, т.е. когда система может, грубо говоря, находиться в одном из двух состояний (см. рис. 7.1). Ясно, что такая система способна хранить информацию, так как одному состоянию мы можем поставить в соответствие 0, а другому 1.

Нормирующий множитель N определяется выражением

$$\int_{-\infty}^{+\infty} N\exp(\alpha\xi^2 - \beta\xi^4)\,d\xi = 1. \tag{7.78}$$

Так как максимумы расположены симметрично относительно начала $\xi = 0$, мы можем заменить левую часть условия (7.78) интегралом

$$2N\int_0^{\infty}\exp(\alpha\xi^2 - \beta\xi^4)\,d\xi. \tag{7.79}$$

При условии, что мы умеем вычислять интеграл в случае, когда имеется только один максимум, интеграл от 0 до ∞ можно заменить интегралом от $-\infty$ до $+\infty$:

$$\int_0^{\infty} \approx \int_{-\infty}^{+\infty}\exp(\alpha\xi^2 - \beta\xi^4)\,d\xi. \tag{7.80}$$

Как показано выше, этот интеграл может быть аппроксимирован величиной

$$\exp\left(\frac{\alpha^2}{4\beta}\right)\int_{-\infty}^{+\infty}\exp(-2\alpha\eta^2)\,d\eta. \tag{7.81}$$

Вычислив интеграл в (7.81), получаем окончательно

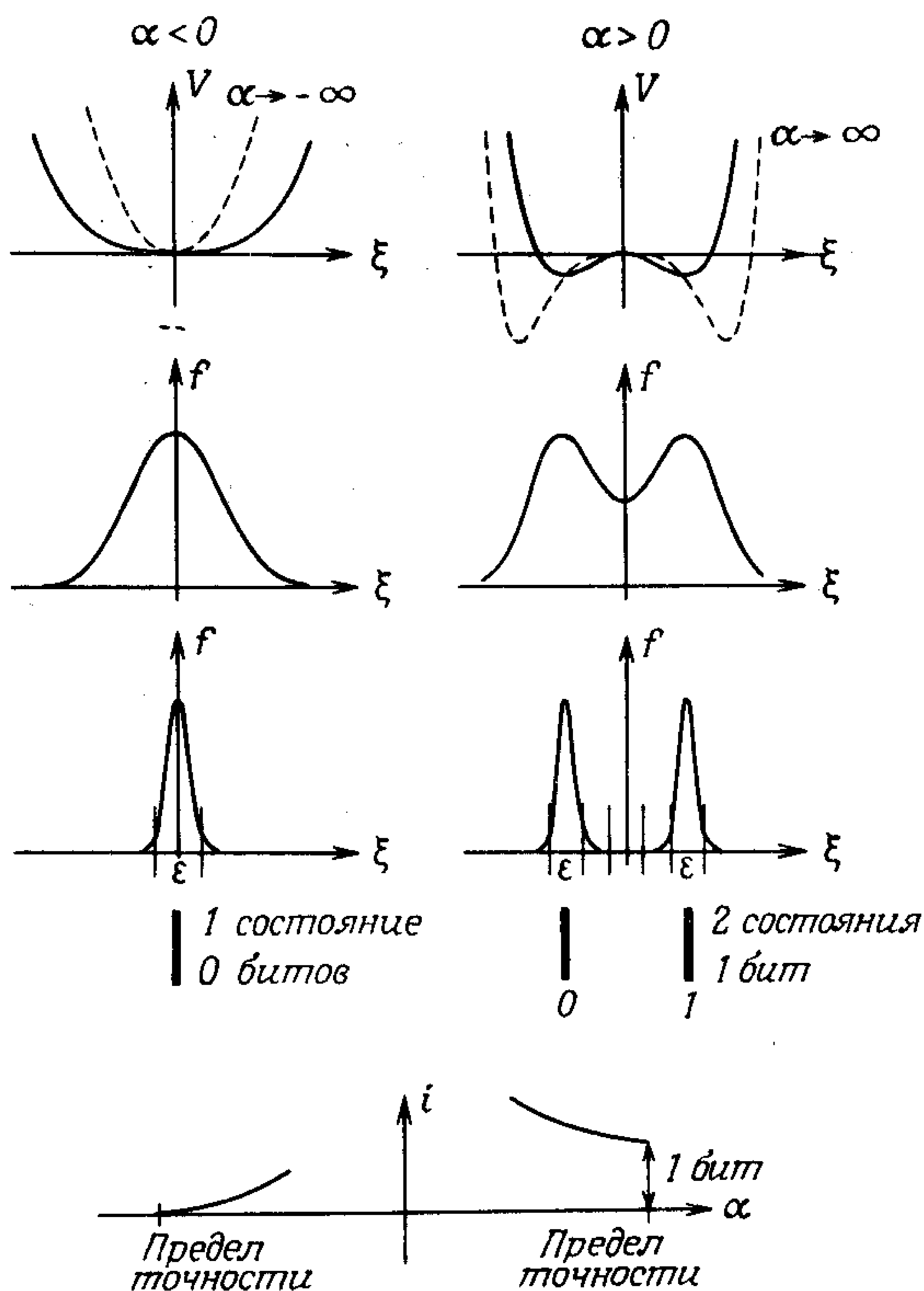


Рис. 7.1. 1-й ряд сверху. Как известно, функцию распределения (7.69) при условии (7.70) можно интерпретировать как функцию распределения для координаты частицы, которая движется в потенциальной яме и подвержена действию дополнительной флуктуирующей силы. Слева показан потенциал при отрицательных α (сплошная линия соответствует значению $|\alpha|$, достаточно близкому к нулю, штриховая – большему значению $|\alpha|$). Совершенно очевидно, что существует только один минимум, к которому может релаксировать частица. Справа показан потенциал при $\alpha > 0$. Сплошная линия соответствует малому положительному значению α, штриховая – большему $\alpha > 0$. Совершенно очевидно, что потенциал стал круче и имеет два минимума.

2-й ряд сверху. Слева: функция распределения f, соответствующая потенциалу с одним минимумом, имеет один максимум. Справа: при $\alpha > 0$ у функции распределения f возникают два различных максимума.

3-й ряд сверху. Слева: при большем значении $|\alpha|$ функция распределения становится более узкой и имеет один максимум. Справа: функция распределения с двумя пиками становится более узкой, и каждая из ее половинок попадает в предел точности ε. Следует подчеркнуть, что масштаб функции f отличается от ее масштаба во 2-м ряду сверху.

4-й ряд сверху. Общий ход зависимости информации от α. Показано, как ведет себя информация вблизи пределов точности.

$$\sqrt{\frac{\pi}{2\alpha}}. \tag{7.82}$$

Соотношения (7.79)–(7.82) позволяют вычислить нормировочный множитель N по формуле (7.78). Он оказывается равным

$$N = \exp\left(-\frac{\alpha^2}{4\beta}\right)\sqrt{\frac{\alpha}{2\pi}}. \tag{7.83}$$

Чтобы вычислить I_f, мы повторяем все этапы вычислений и получаем окончательный результат

$$I_f = -\frac{1}{2}\ln 2\alpha + \frac{1}{2}\ln\pi + \ln 2 + \frac{1}{2} - \ln\varepsilon. \tag{7.84}$$

Сравним соотношение (7.84), которое выполняется в области, расположенной существенно выше порога, с соотношением (7.61), которое выполняется в области, расположенной существенно ниже порога.

Выясним сначала, какую роль играет погрешность измерения ε. В разд. 7.8.1 мы установили соотношение между ε и параметром α, который очевидным образом описывает, как быстро убывают гауссовские функции. Ясно (например, из выражения (7.81)), что постоянная затухания в случае, когда область расположена существенно выше порога, равна 2α, тогда как для области, расположенной существенно ниже порога, та же постоянная равна $|\alpha|$. Исходя из тех же физических соображений, что и раньше, мы должны потребовать выполнения равенства

$$-\frac{1}{2}\ln 2\alpha - \ln\varepsilon = 0. \tag{7.85}$$

Разность между выражениями (7.84) и (7.61) с учетом соотношений (7.85) и (7.67) оказывается равной

$$\Delta I_f = \ln 2. \tag{7.86}$$

Напомним, что для определения информации «по всем правилам» нам следовало выбрать не натуральный логарифм, а логарифм с основанием 2, поэтому

$$\Delta I_f = \log_2 2 \equiv 1. \tag{7.87}$$

Это означает, что система, находящаяся существенно выше порога, может надежно хранить один бит, в то время как система, находящаяся существенно ниже порога, не может хранить информацию. К обсуждению этой интерпретации мы еще вернемся в разд. 7.8.4.

7.8.3. Численные результаты

На рис. 7.2 представлены результаты вычисления I_f при $\varepsilon = 1$ и $\beta = 1$. Относительно результатов разд. 7.8.1 и 7.8.2 интересно заметить, что информация переходит от отрицательных значений параметра α к положительным с образованием максимума. Такую зависимость I_f от α нетрудно понять исходя из физических соображений. Вблизи точки $\alpha = 0$ возникают критические флуктуации, хорошо известные из теории фазовых переходов. Именно они обусловливают расплывание функции f и, следовательно, увеличение I_f. Когда параметр α переходит в область отрицательных значений, функция f оказывается сосредоточенной вокруг одного максимума (см. рис. 7.1) и I_f падает, в то время как при возрастающих положительных значениях α у функции f появляются два различных максимума и информация I_f снова падает, хотя и до более высокого уровня, чем при эквивалентных отрицательных значениях α. Для полноты мы вычислили также приращение информации, определяемое соотношением

$$K = \int_{-\infty}^{+\infty} d\xi f_\alpha \ln \frac{f_\alpha}{f_{\alpha_0}}, \tag{7.88}$$

где под f_α понимается функция

$$f_\alpha = N(\alpha)\exp(\alpha\xi^2 - \beta\xi^4). \tag{7.89}$$

Чтобы получить значения K в окрестности порога, мы положили $\beta = 1$ и вычислили K при фиксированном значении α_0 и параметре α, изменяющемся от -15 до 15 (рис. 7.3–7.5).

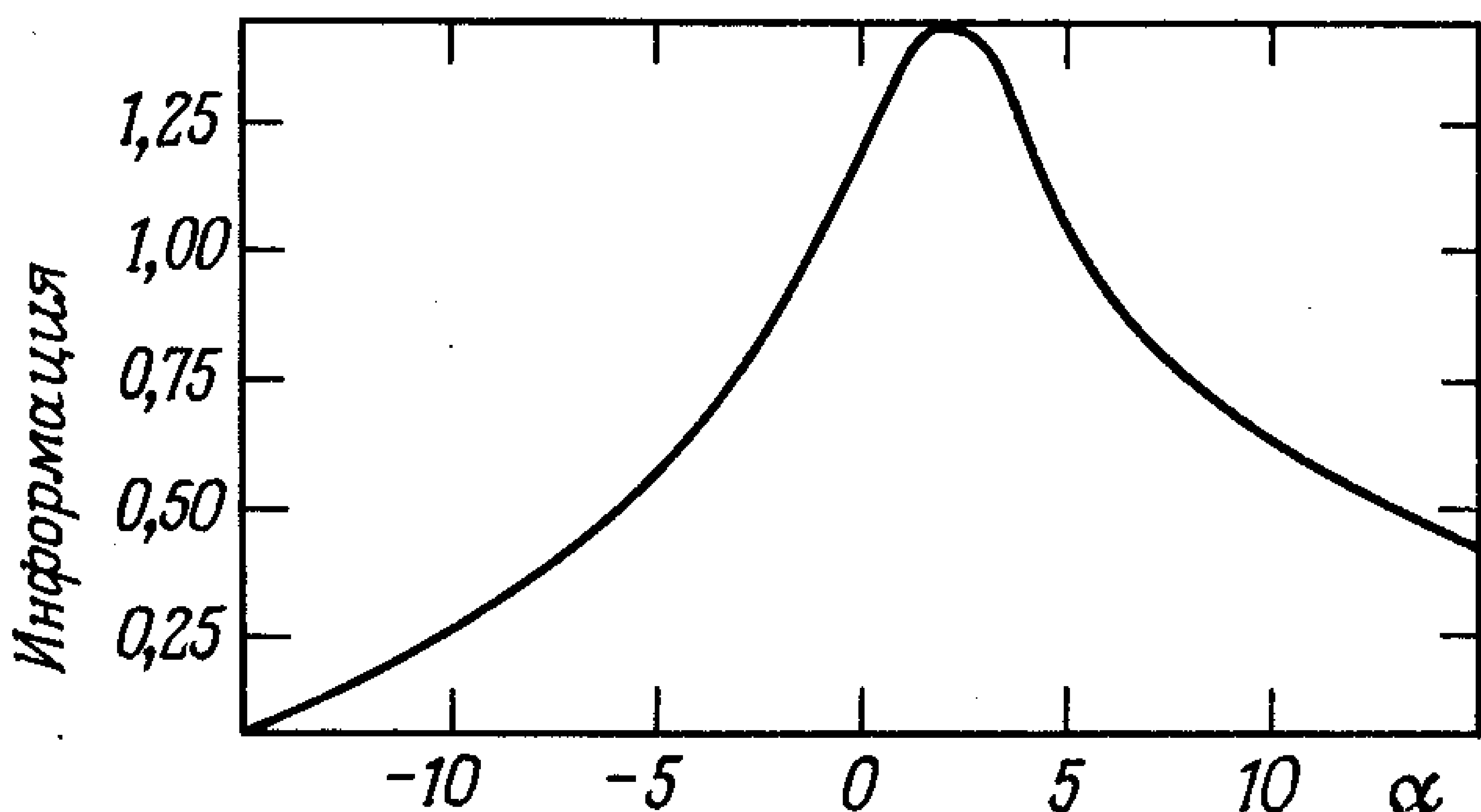


Рис. 7.2. Информация (1.5) как функция от α при $\beta = \varepsilon = 1$.

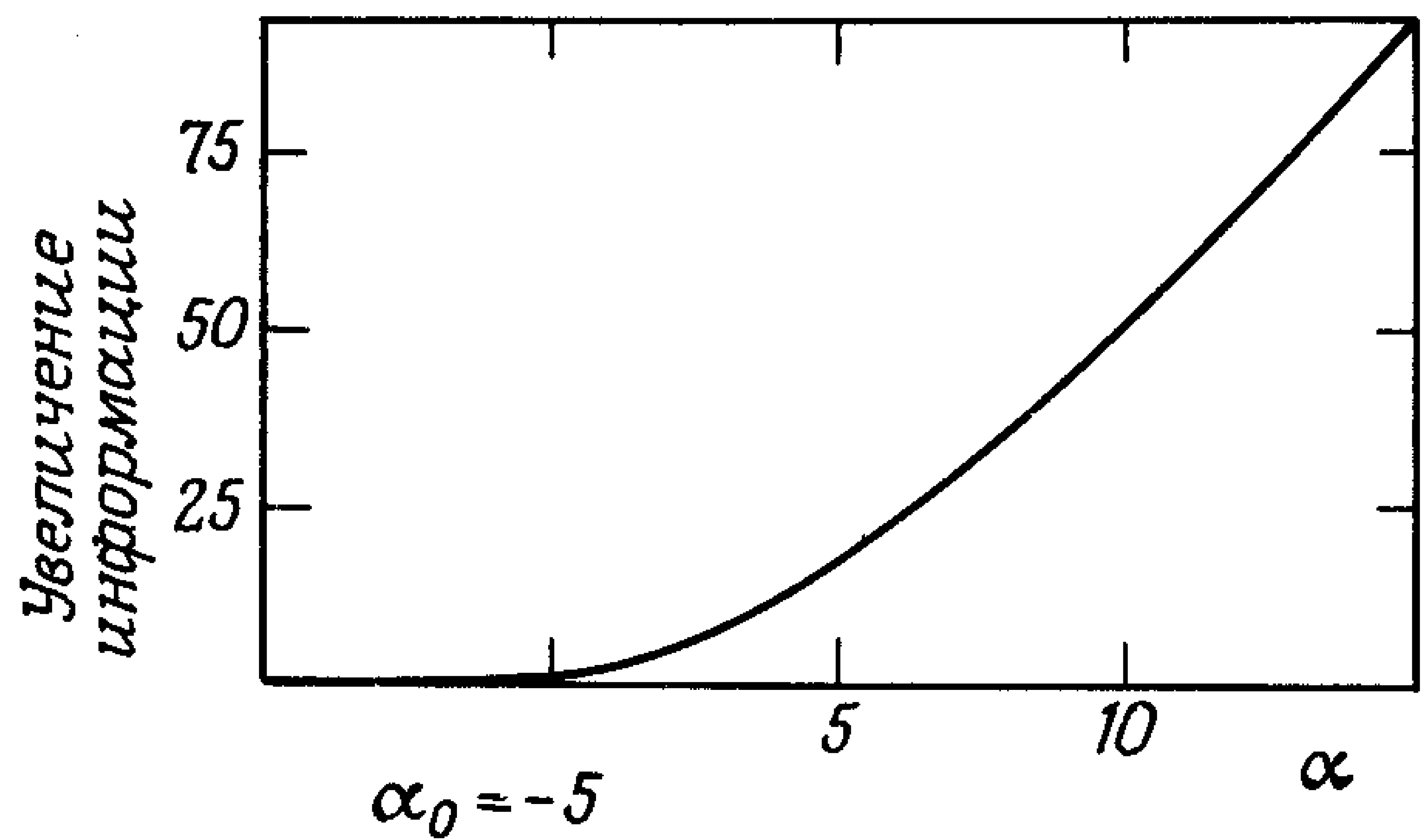


Рис. 7.3. Увеличение информации при $\alpha_0 = -5{,}00$ и коэффициенте α, изменяющемся от -5 до $+15$.

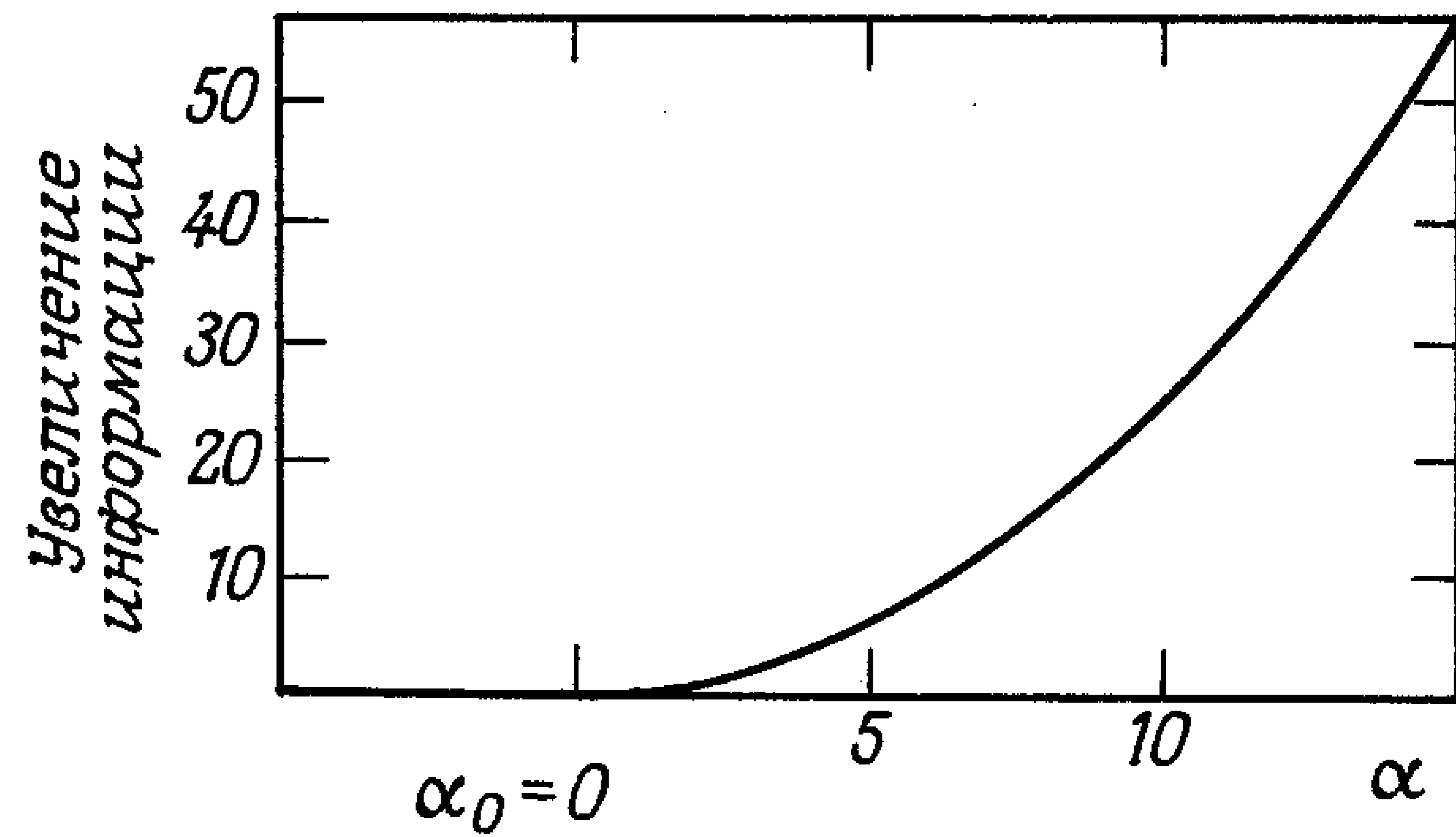


Рис. 7.4. То же, что на рис. 7.3, но при $\alpha_0 = 0$.

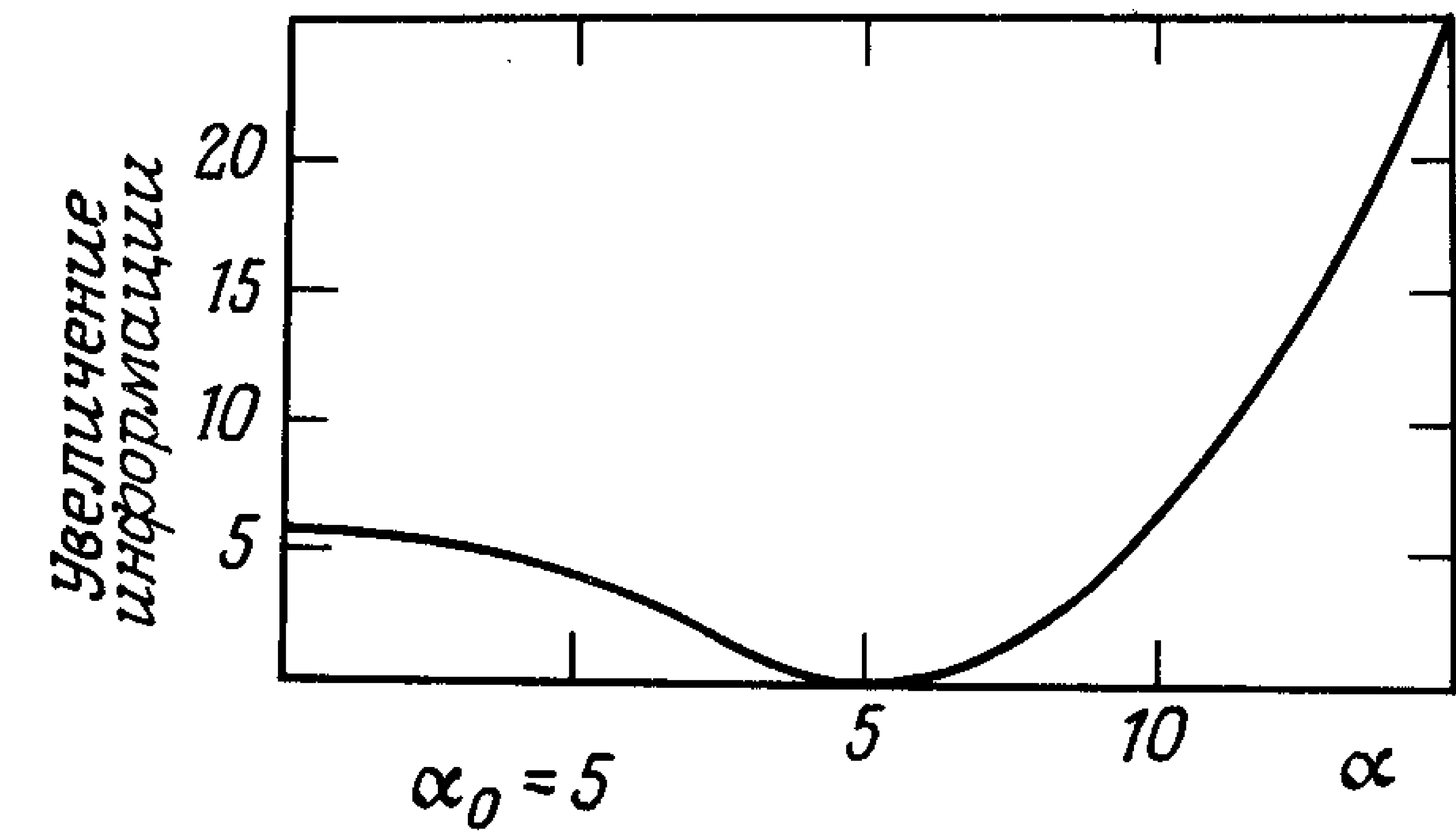


Рис. 7.5. То же, что на рис. 7.3, но при $\alpha_0 = 5$.

7.8.4. Анализ полученных результатов

Мы привели выше подробный ход вычислений изменения информации в самоорганизующейся системе, описываемой одним параметром порядка и претерпевающей фазовый переход второго рода. Как мы уже знаем, изменение информации очень мало, но полностью соответствует тому, чего мы ожидаем исходя из физических и теоретико-информационных соображений. Система обретает возможность хранить один бит информации, если с помощью нарушения симметрии мы можем создавать любое из двух возможных состояний. Это одна из причин, по которым мы отдаем предпочтение понятию информации перед понятием энтропии, когда применяем формулу (7.55) к системам, далеким от теплового равновесия. С другой стороны, поведение вблизи порога показывает, что информация может сильно возрастать здесь из-за критических флуктуаций. В этом случае термин «энтропия» может показаться предпочтительнее. Однако существует еще одна причина, по которой выбор термина «энтропия» вызывает у меня сомнения. Как было показано в гл. 5, ограничения, при которых требуется найти максимум информационной энтропии (7.55), весьма различны для систем, находящихся в тепловом равновесии и далеких от него. В число ограничений, налагаемых на системы, находящиеся в тепловом равновесии, всегда входит ограничение по энергии, в то время как при использовании принципа максимума энтропии соответствующим множителем Лагранжа является обратная температура. В этой связи уместно сделать следующее замечание.

Как хорошо известно из теории неравновесных фазовых переходов, например в лазерах или в гидродинамике, в области $\alpha > 0$ возникают *упорядоченные* структуры. Результаты проведенных нами вычислений, представленные на рис. 7.2, свидетельствуют о том, что в этом упорядоченном состоянии информационная энтропия *выше*, чем в неупорядоченном состоянии. Такой вывод противоречит интуиции, и некоторые авторы утверждают, что для перехода из неупорядоченного состояния в упорядоченное *как в равновесном, так и в неравновесном случае* $dS > 0$. Однако такого рода вывод нельзя считать обоснованным. Ошибочность его станет очевидной, если несколько подробнее рассмотрим термодинамику. Согласно термодинамике, энтропия и, следовательно, беспорядок в *замкнутой* системе могут возрастать, но отнюдь не убывать. Это означает, что если полную энергию системы мы сохраняем постоянной и сравниваем два состояния, то состояние с большей энтропией соответствует большему беспорядку. Но в рассмотренном выше случае *открытой* системы сравнению подлежат два состояния системы, не отвечающие одной и той же энергии. Ясно, что если перекрыть входы и выходы для энергии системы, например лазера, и поддерживать систему при постоянной энергии, то система перестроится, увеличит свою энергию

и перейдет из упорядоченного состояния в некоторое другое состояние. В общем случае мы не можем делать никаких заключений относительно беспорядка или порядка в открытых системах, если рассматриваем только энтропию. С более формальной точки зрения различие в интерпретации энтропии и информационной энтропии обусловлено различием в ограничениях, используемых для замкнутых и для открытых систем.

Предположим, что нам удалось вычислить для *открытой* системы информационную энтропию при соответствующих ограничениях. Вычислим далее среднюю энергию для функции распределения. Если эту энергию принять за новое ограничение, то максимум энтропии достигается *не* на функции распределения открытой системы, а на функции, для которой именно вычисленная средняя энергия является наложенным ограничением. В разд. 7.10 мы поясним этот результат с еще одной точки зрения.

7.9. Точные аналитические результаты относительно информации, приращения информации и эффективности в случае одного параметра порядка

В этом разделе мы приведем точные аналитические выражения для информации I_f, приращения информации K_f и эффективности W. Все эти величины мы вычислим для

$$f(\xi) = N \exp(\alpha\xi^2 - \beta\xi^4). \tag{7.90}$$

Кратко напомним, какими выражениями определяются интересующие нас величины. Информация I_f была определена как величина

$$I_f = -\ln N - \alpha\langle\xi^2\rangle + \beta\langle\xi^4\rangle, \tag{7.91}$$

где

$$\langle\xi^n\rangle = \int\limits_{-\infty}^{+\infty} f(\xi)\,\xi^n\,d\xi \tag{7.92}$$

— n-й момент функции распределения $f(\xi)$. Приращение информации K_f было задано соотношением

$$K_f = K_{ff_0} = \ln N - \ln N_0 + (\alpha - \alpha_0)\langle\xi^2\rangle_\alpha. \tag{7.93}$$

Наконец, так как второй момент $\langle\xi^2\rangle_\alpha$ есть мера интенсивности излучения и, следовательно, произведенной работы, об эффективности системы можно судить по величине

$$W = \frac{d\langle\xi^2\rangle\alpha}{d\alpha}. \tag{7.94}$$

Используя выражение (7.90) для функции распределения, вычисляем

W по формуле (7.94) [см. для сравнения соотношение (7.36)] и получаем

$$W = \langle \xi^4 \rangle_\alpha - \langle \xi^2 \rangle_\alpha^2. \qquad (7.95)$$

Было бы интересно найти соотношение между I_f, K_f, W и управляющими параметрами α и β. Как мы убедились выше, в зависимости от величины параметра α существуют различные режимы, а именно область неустойчивости при $\alpha \approx 0$ и область устойчивости при $|\alpha| \gg 0$.

Основным математическим инструментом, которым мы будем пользоваться, является стандартный интеграл

$$\int_0^\infty x^n \exp(\alpha x^{2m} - \beta x^{4m})\, dx = (2m)^{-1} (2\beta)^{-(n+1)/4m} \Gamma\left(\frac{n+1}{2m}\right) \times$$
$$\times D_{-(n+1)/2m}(\lambda) \exp(\lambda^2/4), \qquad (7.96)$$

где $\lambda = -\alpha/\sqrt{2\beta}$, Γ – гамма-функция и D – функция параболического цилиндра. Так как

$$N^{-1} = \int_{-\infty}^{+\infty} \exp(\alpha\xi^2 - \beta\xi^4)\, d\xi, \qquad (7.97)$$

простое применение формулы (7.96) к интегралу (7.97) дает следующее значение нормировочной постоянной:

$$N = (2\beta)^{1/4} \pi^{-1/2} e^{-\chi^{2/4}} [D_{-1/2}(\lambda)]^{-1}. \qquad (7.98)$$

Аналогичным образом формулы (7.92) и (7.96) позволяют легко вычислить второй и четвертый моменты функции распределения $f(\xi)$:

$$\langle \xi^2 \rangle = (2\beta)^{-1/2} \frac{D_{-3/2}(\lambda)}{2D_{-1/2}(\lambda)}, \qquad (7.99)$$

$$\langle \xi^4 \rangle = (8\beta)^{-1} \frac{3D_{-5/2}(\lambda)}{D_{-1/2}(\lambda)}. \qquad (7.100)$$

Получив выражения, связывающие N, $\langle \xi^2 \rangle$ и $\langle \xi^4 \rangle$ только с управляющими параметрами α и β, мы можем просто подставить их в формулы для I_f, K_f и W, т.е. в формулы (7.91), (7.93) и (7.95). При этом информация, содержащаяся в системе с функцией распределения $f(\xi)$, оказывается равной

$$I_f = 0{,}3991 - \frac{1}{4}\ln\beta + \frac{1}{4}\lambda^2 + \ln D_{-1/2}(\lambda) +$$
$$+ \frac{(\lambda/2) D_{-3/2}(\lambda) + (3/8) D_{-5/2}(\lambda)}{D_{-1/2}(\lambda)}. \qquad (7.101)$$

Аналогичным образом мы вычисляем приращение информации при изменении параметра α, от которого зависит $f(\xi)$, от α_0 до α, и оно оказывается равным

$$K_{ff_0} = \frac{\alpha_0^2 - \alpha^2}{8\beta} + \ln\frac{D_{-1/2}(\lambda_0)}{D_{-1/2}(\lambda)} - (\lambda - \lambda_0)\frac{D_{-3/2}(\lambda)}{2D_{-1/2}(\lambda)}, \qquad (7.102)$$

где $\chi_0 = -\alpha_0/\sqrt{2\beta}$. Наконец, эффективность системы, вычисляемая по формулам (7.95), (7.99) и (7.100), равна

$$W = \frac{3D_{-5/2}(\lambda)D_{-1/2}(\lambda) - [D_{-3/2}(\lambda)]^2}{8\beta[D_{-1/2}(\lambda)]^2}. \qquad (7.103)$$

Следует подчеркнуть, что во все приведенные в этом разделе формулы, а именно соотношения (7.98)–(7.103), входят аналитические функции управляющих параметров α и β, причем $\beta \neq 0$, $\alpha < \infty$. Величины, представляющие для нас особый интерес, а именно I_f, K_f и W, зависят от λ весьма сложным образом. Поскольку D как функция параболического цилиндра принадлежит к числу так называемых специальных функций математической физики, выражения (7.101)–(7.103) в общем случае не поддаются упрощению. Мы можем лишь затабулировать входящие в них функции при отдельных значениях β. Однако существуют три случая, которые заслуживают того, чтобы исследовать их особо. Первый из них возникает, когда точка неустойчивости соответствует $\alpha = 0$ и, следовательно, $\lambda = 0$. Второй случай характеризуется тем, что в непосредственной окрестности точки неустойчивости параметр α и, следовательно, λ стремятся к нулю: $\alpha \to 0$, $\lambda \to 0$. В третьем случае в области устойчивости выполняется неравенство $\alpha \gg 1$. Проанализируем каждый из этих особых случаев подробно.

7.9.1. Точка неустойчивости

Функция параболического цилиндра связана с функцией Вебера соотношением

$$D_{-p-1/2}(\lambda) = U(p, \lambda). \qquad (7.104)$$

В точке неустойчивости $\alpha = 0$ мы используем свойство функции Вебера, состоящее в том, что

$$U(p, 0) = \frac{\sqrt{\pi}}{2^{p/2+1/4}}\Gamma\left(\frac{p}{2} + \frac{3}{4}\right)^{-1}, \qquad (7.105)$$

и обнаруживаем, что нормировочная константа N достигает максимума

$$N_c = 0{,}5516\beta^{1/4}. \qquad (7.106)$$

Второй и четвертый моменты соответственно равны

$$\langle\xi^2\rangle_c = 0{,}3380\beta^{-1/2}, \qquad (7.107)$$

$$\langle\xi^4\rangle_c = 0{,}2550\beta^{-1}. \qquad (7.108)$$

Таким образом, информация I_f в точке неустойчивости определяется выражением

$$I_c = 0{,}8449 - \frac{1}{4}\ln\beta. \qquad (7.109)$$

Аналогично приращение информации при переходе от распределения f_c $(\alpha = 0)$ к распределению f составляет величину

$$K_{ff_c} = 0{,}5949 - \frac{1}{4}\ln\beta + \ln N + \alpha\langle\xi^2\rangle_\alpha. \qquad (7.110)$$

Наконец, эффективность в точке неустойчивости, как нетрудно показать, равна

$$W_c = 0{,}1358\beta^{-1}. \qquad (7.111)$$

Все полученные выражения (кроме выражения, определяющего приращение информации) зависят только от β и, как будет показано ниже, представляют собой максимумы своих функций по α. Интересно исследовать приближение к этим максимумам при $\alpha \to 0$.

7.9.2. *Приближение к точке неустойчивости*

Чтобы исследовать ближайшую окрестность точки неустойчивости, воспользуемся следующим асимптотическим представлением функции Вебера:

$$U(p, \mp|\lambda|) \approx \left[\frac{\sqrt{\pi}}{2^{p/2+1/4}}\Gamma\left(\frac{p}{2}+\frac{3}{4}\right)^{-1}\right]e^{\mp\sqrt{p}\lambda}, \qquad (7.112)$$

где $\lambda \approx 0$ и в разложении экспоненты высшими степенями параметра λ следует пренебречь. Пользуясь асимптотическим представлением (7.112), мы легко получаем главные члены, описывающие поведение величин N, $\langle\xi^2\rangle$ и $\langle\xi^4\rangle$ вблизи $\lambda = 0$:

$$N \approx N_c e^{-\lambda^2}/4, \qquad (7.113)$$

$$\langle\xi^2\rangle \approx \langle\xi^2\rangle_c e^{\mp\lambda}, \qquad (7.114)$$

$$\langle\xi^4\rangle \approx \langle\xi^4\rangle_c e^{\mp\sqrt{2}\lambda}, \qquad (7.115)$$

где N_c, $\langle\xi^2\rangle_c$ и $\langle\xi^4\rangle_c$ определяются выражениями (7.106)–(7.108). В формулах (7.114) и (7.115) знак минус соответствует $\lambda > 0$ $(\alpha < 0)$, а знак плюс – обратному неравенству $\lambda < 0$ $(\alpha > 0)$. Таким образом, при приближении к точке $\alpha = 0$ значения N, $\langle\xi^2\rangle$ и $\langle\xi^4\rangle$ экспоненциально возрастают и стремятся к своим критическим значениям N_c, $\langle\xi^2\rangle_c$ и $\langle\xi^4\rangle_c$. Применяя эти результаты к I_f, получаем

$$I_f \approx I_c + \lambda^2/4 + 0{,}478\lambda\exp(\mp\lambda) + 0{,}250\left[\exp(\mp\sqrt{2}\lambda) - 1\right]. \qquad (7.116)$$

Таким образом, I_f стремится к своему критическому значению I_c и комбинации экспонент от λ. Нетрудно вычислить, что приращение информации, обусловленное изменением функции распределения от f_0 до f_c (и начальная, и конечная функции распределения близки к f_c), равно

$$K_{ff_0} \approx K_{ff_c} + \frac{\alpha_0^2}{8\beta} - \alpha_0 \langle \xi^2 \rangle_\alpha. \tag{7.117}$$

Таким образом, K_{ff_0} стремится к K_{ff_c} через комбинацию членов, один из которых линеен, а другой квадратичен по α_0. Ясно, что $\alpha_0 \to 0$. Полагая $\lambda \approx 0$, мы можем еще больше упростить выражение (7.117) и свести его к следующему:

$$K_{ff_0} \approx \lambda_0^2 - \lambda^2 + 0{,}478\,(\lambda - \lambda_0)\,\mathrm{e}^{\mp\lambda}. \tag{7.118}$$

Наконец, вычисляя эффективность вблизи точки неустойчивости, получаем

$$W \approx W_c \mathrm{e}^{-\sqrt{2}\lambda}. \tag{7.119}$$

В заключение заметим, что все интересующие нас величины, кроме K, стремятся к своим критическим значениям через экспоненты или комбинации экспонент. В этом их поведение резко отличается от поведения в области устойчивости.

7.9.3. Область устойчивости

Другого асимптотического предела, а именно $\lambda \to \infty$, можно достичь, либо увеличивая параметр α, либо фиксируя параметр β, либо устремляя одновременно к бесконечности с одинаковой скоростью параметры α и β, т.е. увеличивая объем системы. Следовательно, асимптотически мы будем отходить от точки неустойчивости. Соответствующее разложение для функции Вебера имеет вид

$$U(p, \lambda) \approx \mathrm{e}^{-\lambda^2/4}\, \lambda^{-p-1/2}, \tag{7.120}$$

где старшие степени параметра $(1/\lambda)$ отброшены. Применяя асимптотическое представление (7.120) к выражению (7.98) для нормировочной постоянной, получаем

$$N \approx (2\beta)^{1/4} \left(\frac{\lambda}{\pi}\right)^{1/2}, \tag{7.121}$$

причем очевидным образом требуется, чтобы $\lambda < 0$. Аналогичным образом для второго момента асимптотическое выражение имеет вид

$$\langle \xi^2 \rangle \approx \frac{1}{2}(2\beta)^{-1/2}\lambda^{-1}, \qquad (7.122)$$

а для четвертого –

$$\langle \xi^4 \rangle \approx \frac{3}{8}\beta^{-1}\lambda^{-2}. \qquad (7.123)$$

Подставляя эти выражения в формулу (7.91), находим асимптотическое выражение для информации

$$I_f = 0{,}5724 - \frac{1}{2}\ln\lambda - \frac{1}{4}\ln(2\beta) + \frac{1}{2} + \frac{3}{8}\lambda^{-2}. \qquad (7.124)$$

Формула I_f также применима лишь при условии, что $\alpha < 0$, и содержит главные члены разложения, описывающего убывание информации I_f до нуля при $\lambda \to \infty$. Разумеется, существует максимальное значение параметра λ, выше которого формула (7.124) становится неприменимой, так как в противном случае она давала бы отрицательную информацию. Кроме того, в подобном случае необходимо также учитывать эффекты, связанные с подчиненными модами, поскольку такие эффекты становятся существенными.

Затем мы вычисляем приращение информации при переходе от f_{α_0} к f_α и получаем

$$K \approx \frac{1}{2}\ln\frac{\alpha}{\alpha_0} - \frac{1}{2}\left(1 - \frac{\alpha_0}{\alpha}\right). \qquad (7.125)$$

Наконец, эффективность системы оказывается равной

$$W \approx \frac{1}{4}\beta^{-1}\lambda^{-2} \qquad (7.126)$$

и монотонно убывает, когда λ стремится к бесконечности. Таким образом, мы видим, что при возрастании λ все интересующие нас величины убывают по различным степенным законам.

7.9.4. Введение в систему сигнала

Без введения внешнего поля, которое связано с ξ через член $-\sigma\xi$, значения ξ с высокой вероятностью двукратно вырождены, т. е. ξ принимает значения $+(\alpha/2\beta)^{1/2}$ и $-(\alpha/2\beta)^{1/2}$ равновероятно. Введение сигнала с амплитудой σ снимает такое вырождение и приводит к тому, что значение ξ того же знака, что и σ, становится более вероятным. Нашу функцию распределения следует при этом заменить следующей:

$$f_\sigma(\xi) = N\exp(-\sigma\xi + \alpha\xi^2 - \beta\xi^4). \qquad (7.127)$$

Условимся в дальнейшем обозначать свойства системы в присутствии

введенного сигнала индексом σ, а в отсутствие сигнала – индексом 0. Располагая функцией распределения $f_\sigma(\xi)$, задаваемой выражением (7.127), мы не можем рассчитывать на получение точных формул, так как к такой функции неприменим стандартный интеграл (7.96). Поэтому мы разложим f_σ в ряд по степеням σ относительно функции f_0, задаваемой формулой (7.90). Так как параметр σ предполагается малым по сравнению с α и β, мы будем учитывать в нашем анализе только члены низшего порядка и аппроксимировать N_σ сначала величиной N_0, задаваемой выражением (7.98). Применяя интеграл (7.96) к разложению функции $f_\sigma(\xi)$ в ряд, получаем аппроксимацию низшего порядка

$$\langle\xi^2\rangle_\sigma \approx \langle\xi^2\rangle_0\left[\frac{1+3}{4\sqrt{2\beta}}\frac{D_{-5/2}(\lambda)}{D_{3/2}(\lambda)}\sigma^2+\ldots\right], \quad (7.128)$$

$$\langle\xi^4\rangle_\sigma \approx \langle\xi^4\rangle_0\left[\frac{1+5}{4\sqrt{2\beta}}\frac{D_{-7/2}(\lambda)}{D_{-5/2}(\lambda)}\sigma^2+\ldots\right], \quad (7.129)$$

где величины $\langle\xi^2\rangle_0$ и $\langle\xi^4\rangle_0$ определяются выражениями (7.99) и (7.100). Амплитуда σ введенного сигнала входит в оба этих выражения через поправочный член, и аппроксимация низшего порядка содержит σ^2. Указав режим изменения λ, мы можем исследовать эти члены более подробно. В окрестности точки неустойчивости получаем

$$\langle\xi^2\rangle_\sigma \approx \langle\xi^2\rangle_0[1+0{,}3698\beta^{-1/2}\exp(\mp 0{,}4142\lambda)\sigma^2], \quad (7.130)$$

$$\langle\xi^4\rangle_\sigma \approx \langle\xi^4\rangle_0[1+0{,}5070\beta^{-1/2}\exp(\mp 0{,}3178\lambda)\sigma^2]. \quad (7.131)$$

Таким образом, поправка на σ умножается на экспоненту от λ и становится все более ощутимой по мере приближения к точке неустойчивости. Мы можем продолжить наши исследования и оценить влияние этой поправки на информацию и эффективность. Они оказываются соответственно равными

$$I_\sigma \approx I_0 - 0{,}1768\beta^{-1/2}\exp(\mp\sqrt{2\lambda})[\lambda - 0{,}7172\exp(\mp 0{,}9036\lambda)]\sigma^2, \quad (7.132)$$

$$W_\sigma \approx W_0 + 0{,}0845\beta^{-1/2}\exp(\mp 2{,}4142\lambda)[1+1{,}5\exp(\pm 0{,}0964\lambda)]\sigma^2. \quad (7.133)$$

Мы видим, что и в этих двух случаях главные члены в поправках квадратичны по σ и умножаются на экспоненты от λ.

С другой стороны, вдали от точки неустойчивости мы получаем для второго и четвертого моментов выражения

$$\langle\xi^2\rangle_\sigma \approx \langle\xi^2\rangle_0(1+0{,}5303\lambda^{-1}\beta^{-1/2}\sigma^2), \quad (7.134)$$

$$\langle\xi^4\rangle_\sigma \approx \langle\xi^4\rangle_0(1+0{,}8839\lambda^{-1}\beta^{-1/2}\sigma^2), \quad (7.135)$$

которые содержат поправки, квадратичные по σ и убывающие линейно по λ^{-1}. Таким образом, информация аппроксимируется выражением

$$I_\sigma \approx I_0 - 0{,}1326\beta^{-1/2}\lambda^{-2}(\alpha - 5\beta\lambda^{-1}), \quad (7.136)$$

а эффективность – выражением

$$W_\sigma \approx W_0 + 0{,}5966\beta^{-1/2}\lambda^{-3}\sigma^2. \quad (7.137)$$

7.9.5. Выводы

В окрестности точки неустойчивости и информация, и эффективность самоорганизующейся системы возрастают экспоненциально как функция от энергии на входе α и, следовательно, с высокой вероятностью подавляют вклад подчиненных мод. С другой стороны, вдали от точки неустойчивости информация и эффективность системы убывают как квадрат энергии на входе α.

Поэтому на сцену в этой области вновь появляются подчиненные моды. Интересно отметить, что величина $\langle\xi^2\rangle^2/\langle\xi^4\rangle$ имеет постоянное значение в обоих пределах. В окрестности точки неустойчивости оно равно 0,4570, что свидетельствует о гораздо большей интенсивности, чем в области устойчивости, где отношение $\langle\xi^2\rangle^2/\langle\xi^4\rangle$ равно только 0,083. Так как мы привели помимо прочего общие формулы для $\langle\xi^2\rangle$, $\langle\xi^4\rangle$, I_f, K_{ff_0} и W, применимые при любых значениях α и β и выражающиеся через аналитические функции, мы ожидаем, что в промежуточной области между $\lambda = \infty$ и $\lambda = 0$ все значения будут заключены в пределах между 0,083 и 0,4750. Введение сигнала приводит к появлению поправочных членов к невозмущенным выражениям. Основной вклад в поправочные члены пропорционален квадрату амплитуды введенного сигнала.

7.10. S-теорема Климонтовича

В этом разделе мы ставим перед собой две задачи. Во-первых, мы хотим показать, что функция распределения для лазера, которая была выведена в разд. 5.2 в рамках нашего макроскопического подхода, может быть получена и в рамках микроскопической теории. Затем мы воспользуемся этим результатом, чтобы изложить одну теорему Климонтовича, которую он назвал S-теоремой и которая позволяет по-новому взглянуть на полученный нами результат, согласно которому в системе, далекой от равновесия, энтропия может возрастать, несмотря на то что система переходит в более упорядоченное состояние.

Но начнем сначала с микроскопической теории. Рассмотрим

комплексный параметр порядка ξ. Нас будет интересовать случай, когда он подчиняется уравнению

$$\dot{\xi} = \lambda\xi - \beta\xi|\xi|^2 + F(t). \tag{7.138}$$

Разложим ξ на вещественную и мнимую части (это понадобится нам в дальнейшем):

$$\xi = q_1 + iq_2. \tag{7.139}$$

Аналогичным образом разложим и флуктуирующую силу F

$$F = F_1 + iF_2, \tag{7.140}$$

где функции F_1 и F_2 предполагаются вещественными. Тогда уравнение (7.138) для параметра порядка также распадается на два уравнения

$$\dot{q}_j = \lambda q_j - \beta q_j(q_1^2 + q_2^2) + F_j, \qquad j = 1, 2, \tag{7.141}$$

для вещественной и мнимой частей q_1 и q_2. Уравнения (7.141) могут быть записаны и кратко:

$$\dot{q}_j = K_j(q_1, q_2) + F_j. \tag{7.142}$$

Как показывает более подробный анализ, флуктуирующие силы обладают следующими свойствами:

$$\langle F_j(t)\rangle = 0, \tag{7.143}$$

$$\langle F_j(t)F_k(t')\rangle = \delta_{jk}Q\delta(t-t'). \tag{7.144}$$

Правила, изложенные в разд. 2.3, позволяют нам сразу же записать уравнение Фоккера–Планка, соответствующее уравнениям (7.141) или (7.138). Оно имеет следующий вид:

$$\dot{f} = -\frac{\partial}{\partial q_1}(K_1 f) - \frac{\partial}{\partial q_2}(K_2 f) + \frac{Q}{2}\left(\frac{\partial^2 f}{\partial q_1^2} + \frac{\partial^2 f}{\partial q_2^2}\right). \tag{7.145}$$

Нетрудно убедиться в том, что условия детального равновесия выполнены. Таким образом, для стационарного состояния, когда

$$\dot{f} = 0, \tag{7.146}$$

получаем

$$f(q_1, q_2) = \mathcal{N}\exp\{[\lambda(q_1^2 + q_2^2) - \frac{\beta}{2}(q_1^2 + q_2^2)^2]Q^{-1}\}. \tag{7.147}$$

Для дальнейшего нам удобно произвести замену переменных в функции распределения. Для этого рассмотрим вероятность

$$f(q_1, q_2)\,dq_1\,dq_2. \tag{7.148}$$

Введем новые переменные по формулам

$$q_1 = r\cos\phi, \tag{7.149}$$

$$q_2 = r\sin\phi, \tag{7.150}$$

или, что эквивалентно,

$$\xi = r\exp(i\phi). \tag{7.151}$$

В новых переменных r и ϕ элемент объема может быть записан в виде

$$dV = dq_1\,dq_2 = r\,dr\,d\phi. \tag{7.152}$$

Таким образом, вероятность (7.148) может быть записана в переменных r и ϕ следующим образом:

$$f(r)\,r\,dr\,d\phi = \mathcal{N}\exp\left[\left(\lambda r^2 - \frac{1}{2}\beta r^4\right)Q^{-1}\right]r\,dr\,d\phi. \tag{7.153}$$

Введем теперь величину U, такую, что

$$r^2 = U. \tag{7.154}$$

В ряде физических задач эту величину можно интерпретировать как потенциальную энергию. Из соотношений (7.153) и (7.154) получаем

$$f(U)\,dU = \mathcal{N}\exp\left[\left(\lambda U - \frac{1}{2}\beta U^2\right)Q^{-1}\right]dU, \tag{7.155}$$

где мы опустили справа и слева общий множитель $d\phi$. В дальнейшем нам понадобится следующая нормировка функции $f(U)$:

$$\int_0^\infty f(U)\,dU = 1. \tag{7.156}$$

Чтобы подготовить все необходимое для объяснения S-теоремы, рассмотрим информационную энтропию

$$i = -\int_0^\infty f(U)\ln f(U)\,dU \tag{7.157}$$

и среднюю энергию, определяемую соотношением

$$\langle U\rangle = \int_0^\infty UF(U)\,dU. \tag{7.158}$$

Чтобы записать подробно формулы (7.157) и (7.158), проанализируем три предельных случая, хорошо известные из лазерной физики, но могущие представиться и в других областях.

7.10.1. Область 1: значительно ниже порога генерации

В этой области справедливо неравенство $\lambda < 0$. Лазерный генератор колебаний возбуждается лишь слабо, и его энергия удовлетворяет неравенству

$$\frac{\beta}{2}U^2 \ll |\lambda|U. \tag{7.159}$$

При таком допущении мы можем пренебречь квадратичным членом по сравнению с линейным в функции распределения (7.147), что позволяет просто вычислить среднюю энергию. Она оказывается равной

$$\langle U \rangle = \frac{Q}{|\lambda|}. \tag{7.160}$$

Подставляя выражение (7.160) в неравенство (7.159), мы можем записать последнее в виде

$$\frac{\beta Q}{2|\lambda|^2} \ll 1. \tag{7.161}$$

Используемая в этом случае функция распределения определяется соотношением

$$f_1(U) = \frac{|\lambda|}{Q}\exp\left(-\frac{|\lambda|U}{Q}\right), \tag{7.162}$$

в которое мы включили точный нормирующий множитель. Вычисление величин (7.157) и (7.158) не составляет особого труда, и мы получаем

$$i_1 = \ln\frac{Q}{|\lambda|} + 1, \tag{7.163}$$

$$\langle U \rangle = \frac{Q}{|\lambda|}. \tag{7.164}$$

Поскольку мы вычисляем величины (7.157) и (7.158) в режиме 1, т.е. ниже порога генерации, величинам i и $\langle U \rangle$ также необходимо приписать индекс 1. При увеличении накачки лазера, т.е. при повышении мощности накачки, параметр λ возрастает. Соотношение (7.164) показывает, что при этом средняя энергия также увеличивается. Рассмотрим теперь другую область, в которой лазер находится на пороге генерации.

7.10.2. Область 2: на пороге генерации

В этом случае управляющий параметр λ удовлетворяет условию

$$\lambda = 0. \tag{7.165}$$

Соответствующая функция распределения имеет вид

$$f_2(U) = \sqrt{\frac{2\beta}{\pi Q}} \exp\left(-\frac{\beta U^2}{2Q}\right). \tag{7.166}$$

Индекс 2 указывает на то, что система находится на пороге генерации. И в этом случае информационная энтропия и средняя энергия вычисляются без труда:

$$i_2 = \frac{1}{2}\ln\frac{\pi Q}{2\beta} + \frac{1}{2}, \tag{7.167}$$

$$\langle U\rangle_2 = \sqrt{\frac{2Q}{\pi\beta}}. \tag{7.168}$$

7.10.3. Область 3: значительно выше порога генерации

В этой области справедливо неравенство $\lambda > 0$. Нас будет интересовать подобласть, в которой выполняется неравенство

$$\frac{Q\beta}{\lambda^2} \ll 1. \tag{7.169}$$

Функция распределения f в области (7.169) хорошо аппроксимируется выражением

$$f_3(U) = \mathcal{N} \exp\left[-\frac{\beta}{2Q}\left(U - \frac{\lambda}{\beta}\right)^2\right], \tag{7.170}$$

где индекс 3 указывает на то, что мы рассматриваем область, расположенную далеко за порогом генерации. Информационная энтропия и средняя энергия могут быть вычислены без труда, и мы получаем

$$i_3 = \frac{1}{2}\ln\frac{2\pi Q}{\beta} + \frac{1}{2}, \tag{7.171}$$

$$\langle U\rangle_3 = \frac{\lambda}{\beta}. \tag{7.172}$$

Сравним информационные энтропии и средние энергии в областях 1, 2 и 3. Нетрудно убедиться, что они удовлетворяют неравенствам

$$i_1 < i_2 < i_3, \tag{7.173}$$

$$\langle U\rangle_1 < \langle U\rangle_2 < \langle U\rangle_3. \tag{7.174}$$

С увеличением накачки информационная энтропия возрастает, как это уже было обнаружено нами в разд. 7.9 для вещественного параметра порядка. В том же разд. 7.9 мы заметили, что адекватное сравнение энтропий возможно только при одинаковых энергиях. Поэтому, следуя идее Климонтовича, мы выбираем интенсивность

флуктуаций Q так, чтобы энергии были равны. Иначе говоря, мы изменяем флуктуирующие силы в областях 1 и 2 так, чтобы выполнялись равенства

$$\langle U \rangle_{1,r} = \langle U \rangle_{2,r} = \langle U \rangle_{3,r}. \tag{7.175}$$

Мы хотим подробно исследовать, к каким следствиям приводит сравнение состояний при одной и той же энергии, и начнем с области, лежащей ниже порога генерации.

Чтобы сохранить постоянную энергию при возрастающем параметре λ, необходимо увеличить уровень шума Q соответственно величине λ. Интересно обсудить возникающую ситуацию с точки зрения лазерной физики. При уменьшении параметра λ, т. е. при увеличении интенсивности накачки, возникает процесс индуцированного излучения. Это означает, что отдельные цуги волн поля излучения усиливаются. В результате средняя энергия возрастает. Вместе с тем при понижении уровня шума становятся более редкими акты спонтанного излучения. Если понижать уровень шума или скорость производства спонтанного излучения в соответствии с коэффициентом усиления, то нетрудно найти условия, при которых энергия остается постоянной. В то же время мы обнаруживаем, что информационная энтропия остается постоянной, т. е. в так называемой области линейного усиления информационная энтропия остается постоянной при описанных выше условиях.

Климонтович выбрал другое сравнение. Он воспользовался равенствами (7.175) при фиксированной средней энергии, превышающей порог генерации:

$$\langle U \rangle_{3,r} = \frac{\lambda}{\beta}, \tag{7.176}$$

где λ и β – фиксированные величины. В соотношении

$$\langle U \rangle_{2,r} = \sqrt{\frac{2Q}{\pi\beta}} \tag{7.177}$$

мы полагаем $Q = Q_2$. Тогда соотношение

$$\langle U \rangle_{2,r} = \frac{\lambda}{\beta} \tag{7.178}$$

выполняется, а для Q_2 получаем

$$Q_2 = \frac{\pi}{2}\frac{\lambda^2}{\beta}. \tag{7.179}$$

Подставляя правую часть соотношения (7.179) в выражение для информационной энтропии (7.167), получаем

$$i_{2,r} = \frac{1}{2}\ln\frac{\pi Q_2}{2\beta} + \frac{1}{2} = \ln\frac{\pi\lambda}{2\beta} + \frac{1}{2}. \tag{7.180}$$

Проделаем то же самое в области 1. Пусть в соотношении

$$\langle U \rangle_{1,r} = \frac{Q}{|\lambda_1|} \tag{7.181}$$

$Q = Q_1$. Тогда

$$\langle U \rangle_{1,r} = \frac{\lambda}{\beta}. \tag{7.182}$$

Нетрудно видеть, что

$$Q_1 = \frac{\lambda|\lambda_1|}{\beta}. \tag{7.183}$$

Подставляя этот результат в выражение для информационной энтропии (7.163), находим

$$i_{1,r} = \ln\frac{Q_1}{|\lambda_1|} + 1 = \ln\frac{\lambda}{\beta} + 1. \tag{7.184}$$

Вспоминая неравенство

$$\frac{Q\beta}{\lambda^2} \ll 1, \tag{7.185}$$

приходим к следующим неравенствам для энтропий:

$$i_{1,r} > i_{2,r} > i_{3,r}. \tag{7.186}$$

Таким образом, если среднюю энергию поддерживать постоянной за счет адекватного выбора уровня шума, то энтропия действительно убывает с увеличением управляющего параметра и возникает более упорядоченное состояние системы, например генерация лазерного излучения. Утверждение о том, что информационная энтропия при переходе в более упорядоченное состояние убывает, т.е. о том, что происходит явление самоорганизации, было названо Климонтовичем S-теоремой. Этот же автор рассмотрел также общий случай областей, занимающих промежуточное положение между нашими областями 1 и 3. Свою теорему Климонтович доказал, рассматривая изменение энергии и информационной энтропии при изменении управляющего параметра.

В первоначальном варианте доказательства S-теоремы Климонтович рассматривал только изменение энергии, обусловленное управляющим параметром, но не учитывал одновременного изменения уровня шума, необходимое для сохранения постоянства энергии. Следовательно, первоначальное доказательство S-теоремы требует

небольших изменений, на которых, однако, мы не будем здесь останавливаться.

7.11. Вклад подчиненных мод в информацию вблизи неравновесных фазовых переходов

В разд. 7.2 мы показали, что информация всей системы допускает разложение на информацию параметров порядка и информацию подчиненных мод. В предыдущих разделах мы вычислили информацию параметров порядка. Обратимся теперь к исследованию информации подчиненных мод. Воспользуемся для этого принципом подчинения, точнее главными членами его аппроксимации. Как было показано в соотношении (7.21), условная вероятность P при этом определяется соотношением

$$P(\xi_s | \xi_u) = \mathcal{N}_s \exp\left\{-\frac{[\lambda_s \xi_s + g_s(\xi_u)]}{Q_s}\right\}, \tag{7.187}$$

где

$$\mathcal{N}_s = |\lambda_s| (\pi Q_s)^{-1/2} \tag{7.188}$$

–нормирующий множитель. В дальнейшем мы будем считать параметры λ_s, Q_s и, следовательно, нормирующий множитель $\mathcal{N}_s$ функциями управляющего параметра α:

$$\lambda_s = \lambda_s(\alpha), \tag{7.189}$$

$$Q_s = Q_s(\alpha), \tag{7.190}$$

$$\mathcal{N}_s = \mathcal{N}_s(\alpha). \tag{7.191}$$

Кроме того, мы можем предположить, что g_s зависит от этого управляющего параметра α и что

$$g_s(0) = 0. \tag{7.192}$$

Информация моды s при данном фиксированном значении управляющего параметра ξ_u определяется соотношением

$$I_s(\xi_u, \alpha) = -\int P(\xi_s | \xi_u) \ln P(\xi_s | \xi_u) d\xi_s. \tag{7.193}$$

Интеграл в правой части (7.193) берется от минус бесконечности до плюс бесконечности. Введем вместо ξ_s новую переменную ξ', удовлетворяющую соотношению

$$\xi_s + \frac{g_s}{\lambda_s} = \xi'. \tag{7.194}$$

Подставляя соотношения (7.187) и (7.194) в выражение (7.193), получаем

$$I_s = -\int \mathcal{N}_s(\alpha)\exp\left(-\frac{\lambda_s^2 \xi'^2}{Q_s}\right)\cdot \ln\left[\mathcal{N}_s(\alpha)\exp\left(-\frac{\lambda_s^2 \xi'^2}{Q_s}\right)\right]d\xi'. \quad (7.195)$$

Чтобы исключить, насколько это возможно, зависимость I_s от α, произведем подстановку

$$|\lambda_s(\alpha)|\,Q_s(\alpha)^{-1/2}\,\xi' = |\lambda_s(0)|\,Q_s(0)^{-1/2}\,\eta, \quad (7.196)$$

или, если разрешить равенство (7.196) относительно ξ',

$$\xi' = \frac{|\lambda_s(0)|}{|\lambda_s(\alpha)|}\left(\frac{Q_s(0)}{Q_s(\alpha)}\right)^{-1/2}\eta. \quad (7.197)$$

Тогда рассматриваемая как функция параметра порядка ξ_u и управляющего параметра α информация I_s представима в виде

$$I_s(0, \alpha) = -\int \frac{\mathcal{N}_s(\alpha)}{\mathcal{N}_s(0)}\mathcal{N}_s(0)\exp\left[-\frac{\lambda_s(0)^2\eta^2}{Q_s(0)}\right]\times$$
$$\times \ln\left\{\frac{\mathcal{N}_s(\alpha)}{\mathcal{N}_s(0)}\mathcal{N}_s(0)\exp\left[-\frac{\lambda_s(0)^2\eta^2}{Q_s(0)}\right]\right\}\times$$
$$\times \frac{|\lambda_s(0)|}{|\lambda_s(\alpha)|}\left(\frac{Q_s(0)}{Q_s(\alpha)}\right)^{-1/2}d\eta. \quad (7.198)$$

Следует подчеркнуть, что выражение (7.198) не зависит от параметра порядка ξ_u. На это указывает аргумент 0 в I_s. Анализируя соотношение (7.198), нетрудно заметить, что несколько множителей в нем сокращается и оно приводится к более простой форме

$$I_s(0, \alpha) = -\mathcal{N}_s(0)\int \exp\left[-\frac{\lambda_s(0)^2\eta^2}{Q_s(0)}\right]\{A + B\}\,d\eta, \quad (7.199)$$

где A определяется по формуле

$$A = \ln\frac{\mathcal{N}_s(\alpha)}{\mathcal{N}_s(0)}, \quad (7.200)$$

а B – по формуле

$$B = \ln\left\{\mathcal{N}_s(0)\exp\left[-\frac{\lambda_s(0)^2\eta^2}{Q_s(0)}\right]\right\}. \quad (7.201)$$

Так как A – константа, интеграл от гауссовской функции может быть вычислен без труда и при надлежащем нормирующем множителе оказывается равным единице. Нетрудно убедиться в том, что интеграл, содержащий член B, есть не что иное, как определение информации I_s при управляющем параметре $\alpha = 0$. Следовательно, соотношение (7.199) представимо в виде

$$I_s(0, \alpha) = -\ln \frac{\mathcal{N}_s(\alpha)}{\mathcal{N}_s(0)} + I_s(0, 0). \tag{7.202}$$

Используя выражение (7.188) для нормирующего множителя $\mathcal{N}_s$, соотношение (7.202) можно записать в форме

$$I_s(0, \alpha) = -\ln\left[\frac{|\lambda_s(\alpha)|}{|\lambda_s(0)|}\left(\frac{Q_s(\alpha)}{Q_s(0)}\right)^{-1/2}\right] + I_s(0, 0). \tag{7.203}$$

Напомним теперь общее разложение информации i всей системы на информацию параметров порядка I_f и информацию подчиненных мод. Так как информация I_s не зависит от ξ_u, интегрирование по параметрам порядка выполняется без труда и приводит к множителю 1. Окончательно получаем

$$i = I_f + \sum_s I_s(0, 0) - \sum_s \ln \frac{\mathcal{N}_s(\alpha)}{\mathcal{N}_s(0)}. \tag{7.204}$$

Если требуется сравнить информацию при двух различных значениях α_2 и α_1 управляющего параметра, то образуем разность

$$i(\alpha_2) - i(\alpha_1) = I_u(\alpha_2) - I_u(\alpha_1) - \sum_s \ln \frac{\mathcal{N}_s(\alpha_2)}{\mathcal{N}_s(\alpha_1)}. \tag{7.205}$$

Заметим, что $\mathcal{N}^{-1}$ соответствует известной из термодинамики функции распределения, а $\ln \mathcal{N}$, как следует из соотношения

$$\ln \mathcal{N} = -\lambda, \tag{7.206}$$

соответствует свободной энергии. Ясно, что формула (7.204) позволяет обобщить понятия термодинамики. Если управляющий параметр α выбран так, что в критической точке выполняется соотношение $\alpha = 0$, то, взяв производную, мы без труда вычислим изменение информации I_s в окрестности критической точки. Оно оказывается равным

$$\left.\frac{dI_s}{d\alpha}\right|_{\alpha=0} = -\frac{\mathcal{N}'_s(0)}{\mathcal{N}_s(0)}. \tag{7.207}$$

Пример: одномодовый лазер

Исследуем теперь, как ведет себя информация подчиненных мод вблизи неравновесных фазовых переходов на конкретном примере. Одним из простейших примеров такого рода может служить одномодовый лазер. В этой связи мы хотим кратко напомнить читателю соответствующую теорию. Лазер описывается моделью одномодового поля с амплитудой $b(t)$, которую записывают в безразмерной форме. Величина $|b|^2$ может быть интерпретирована как среднее число фотонов в лазере.

Активная среда лазера состоит из атомов, обладающих по предположению только двумя уровнями. Атомный дипольный момент описывается безразмерной переменной α_μ. Наконец, инверсия для каждого атома, т.е. разность населенностей верхнего и нижнего уровней, определяется величиной d_μ. Уравнения лазера принимают при этом следующий вид:

$$\dot{b} = (-\mathrm{i}\omega - \varkappa) b - \mathrm{i}g \sum_\mu \alpha_\mu, \tag{7.208}$$

$$\dot{\alpha}_\mu = (-\mathrm{i}\omega - \gamma)\alpha_\mu + \mathrm{i}g\, b d_\mu, \tag{7.209}$$

$$\dot{d}_\mu = \gamma_\parallel (d_0 - d_\mu) + 2\mathrm{i}g(\alpha_\mu b^* - \alpha_\mu^* b), \tag{7.210}$$

где ω – частота оптического перехода атомов активной среды лазера, $\varkappa$ – коэффициент затухания полевой моды, g – константа связи между полевой модой и атомным дипольным моментом γ, $\gamma_\parallel$ – постоянная релаксации инверсной населенности атома, d_0 – заданная инверсия, создаваемая интенсивностью накачки, с которой возбуждаются атомы, и процессами некогерентной релаксации.

Чтобы исключить постоянную ω, мы примем гипотезу, согласно которой

$$b = B\mathrm{e}^{-\mathrm{i}\omega t}, \qquad \alpha_\mu = A_\mu \mathrm{e}^{-\mathrm{i}\omega t}, \tag{7.211}$$

где B и A_μ – зависящие от времени переменные. Подставляя зависимости (7.211) в уравнения лазера (7.208) – (7.210), получаем

$$\dot{B} = -\varkappa B - \mathrm{i}g \sum_\mu A_\mu, \tag{7.212}$$

$$\dot{A}_\mu = -\gamma A_\mu + \mathrm{i}g\, b d_\mu, \tag{7.213}$$

$$\dot{d}_\mu = \gamma_\parallel (d_0 - d_\mu) + 2\mathrm{i}g(A_\mu B^* - A_\mu^* B). \tag{7.214}$$

Величину d_0 можно идентифицировать как управляющий параметр. Как известно, при слабой инверсии d_0 генерации лазерного излучения не происходит. Рассмотрим сначала именно этот случай. Поле исчезает, дипольные моменты атомов равны нулю, в то время как инверсность d_μ принимает значение, равное значению d_0, создаваемому накачкой и процессами некогерентной релаксации. Таким образом, выполняются следующие равенства:

$$B_0 = 0, \qquad A_{\mu,0} = 0 \qquad d_{\mu,0} = d_0. \tag{7.215}$$

Мы хотим выяснить, при каких значениях d_0 решение (7.215) устойчиво и при каких критических значениях d_0 оно теряет устойчивость. Для этого мы проделываем обычный анализ устойчивости по линейному приближению, описанный в разд. 2.6, а именно предполагаем, что

$$B = B_0 + \delta B, \qquad A_\mu = A_{\mu,0} + \delta A_\mu. \tag{7.216}$$

Подставляя разложения (7.216) в уравнения (7.212)–(7.214) и используя равенства (7.215), получаем в качестве главных, т. е. линейных, членов аппроксимации уравнения

$$\delta \dot{B} = -\varkappa \delta B - \mathrm{i} g \sum_{\mu} \delta A_{\mu}, \qquad (7.217)$$

$$\delta \dot{A}_{\mu} = -\gamma \delta A_{\mu} + \mathrm{i} g \delta B d_0. \qquad (7.218)$$

Кроме того оказывается, что уравнение (7.214) не дает изменения параметра d_{μ} в первом порядке. Заметим, что уравнения (7.217) и (7.218) имеют такой же вид, как уравнение (2.72) из разд. 2.6, и могут служить примером выведенной там системы линейных уравнений. Для решения уравнений (7.217) и (7.218) предположим, как обычно, что

$$\delta B(t) = \delta B(0) \mathrm{e}^{\lambda t}, \qquad (7.219)$$

$$\delta A_{\mu}(t) = \delta A_{\mu}(0) \mathrm{e}^{\lambda t}, \qquad (7.220)$$

где не зависящие от времени амплитуды и параметр λ – пока не известные величины. Подставляя выражения (7.219) и (7.220) в уравнения (7.217) и (7.218), получаем линейные алгебраические уравнения

$$\delta B(0)(\lambda + \varkappa) = -\mathrm{i} g \sum_{\mu} \delta A_{\mu}(0), \qquad (7.221)$$

$$\delta A_{\mu}(0)(\lambda + \gamma) = \mathrm{i} g \delta B(0) d_0. \qquad (7.222)$$

Определитель системы уравнений (7.221), (7.222) имеет следующий вид:

$$\begin{vmatrix} -\varkappa - \lambda & -\mathrm{i}g & -\mathrm{i}g & \ldots & -\mathrm{i}g \\ \mathrm{i}gd_0 & -\gamma - \lambda & 0 & \ldots & 0 \\ \mathrm{i}gd_0 & 0 & -\gamma - \lambda & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathrm{i}gd_0 & 0 & 0 & \ldots & -\gamma - \lambda \end{vmatrix}$$

Его нетрудно вычислить, и мы получаем характеристическое уравнение

$$-(\varkappa + \gamma)(\gamma + \lambda)^N + g^2 d_0 N (\gamma + \lambda)^{N-1} = 0. \qquad (7.223)$$

Собственные значения λ имеют следующие значения:

а) для устойчивых мод

$$\lambda_1 = -\frac{\varkappa + \gamma}{2} - \frac{1}{2}\sqrt{(\varkappa - \gamma)^2 + 4g^2 D_0};$$

$$D_0 = N d_0 \qquad (\lambda_1 \text{ – простое собственное значение}) \qquad (7.224)$$

и при $s = 2, \ldots, N$

$\lambda_s = -\gamma$ (λ_s–$(N-1)$-кратно вырожденное собственное значение); (7.225)

б) для неустойчивой моды

$$\lambda_u = -\frac{\varkappa+\gamma}{2} + \frac{1}{2}\sqrt{(\varkappa-\gamma)^2 + 4g^2 D_0} \tag{7.226}$$

или, в другой форме,

$$\lambda = -\frac{\varkappa+\gamma}{2} + \frac{1}{2}\sqrt{(\varkappa-\gamma)^2 + 4g^2 D_0}. \tag{7.227}$$

Критическое значение инверсности D_c следует из условия неустойчивости и оказывается равным

$$\lambda_u \geqslant 0; \qquad D_c = \frac{\varkappa\gamma}{g^2}. \tag{7.228}$$

Теперь мы хотим исследовать изменение информации подчиненных мод вблизи точки неустойчивости. Положим для этого

$$D_0 = D_c + \alpha' = \frac{\varkappa\gamma}{g^2} + \alpha', \tag{7.229}$$

где параметр α' играет роль управляющего параметра.

Следуя обычной теории лазера, предположим теперь, что флуктуирующие силы не зависят от управляющего параметра. Если принять это предположение и учесть соотношение (7.188), то формула (7.207) сводится к следующей:

$$\left.\frac{dI_s}{d\alpha'}\right|_{\alpha'=0} = -\frac{\lambda_s'(\alpha'=0)}{\lambda_s(\alpha'=0)}, \qquad s = 1, \ldots, N. \tag{7.230}$$

Нетрудно показать, что при $s = 2, \ldots, N$ производная (7.230) равна нулю и что

$$\frac{\lambda_1'(0)}{\lambda_1(0)} = \frac{g^2}{(\varkappa+\gamma)^2} \approx \frac{g^2}{\gamma^2} = \frac{\varkappa}{\gamma}\frac{1}{D_c}. \tag{7.231}$$

Формула (7.231) получена для частного случая, когда

$$\varkappa \ll \gamma. \tag{7.232}$$

Неравенство (7.232) справедливо для большинства лазеров. Таким образом, наш окончательный результат может быть записан следующим образом:

$$\frac{dI_s}{d\alpha'}\alpha' = \begin{cases} -\dfrac{\varkappa}{\gamma}\dfrac{\alpha'}{D_c}, & \text{если} \quad s = 1; \\ 0, & \text{если} \quad s \neq 1. \end{cases} \tag{7.233}$$

Он свидетельствует о том, что с увеличением скорости накачки, характеризуемой параметром α', информация убывает. Этот вывод нетрудно интерпретировать, если вспомнить, что общая форма распределения вероятности подчиненных мод определяется выражением (7.187). Если Q_s – константа, а величина g_s оттрансформирована по формуле (7.194), то информация подчиненных мод зависит от гауссовской функции распределения, ширина которой убывает при увеличении абсолютной величины параметра λ_s (что соответствует увеличению параметра α'). Однако более узкая функция распределения означает бóльшую определенность и в то же время свидетельствует об уменьшении информации.

Обращаясь к аппроксимации принципа подчинения на уровне главных членов, мы замечаем, что информационная энтропия атомов лазера убывает при более интенсивной накачке лазера. Наконец, свяжем управляющий параметр α', введенный с помощью соотношения (7.229), с управляющим параметром α, входящим в функцию распределения параметра порядка

$$f = \mathcal{N} \exp(\alpha\xi^2 - \beta\xi^4). \tag{7.234}$$

Для этого нам понадобится напомнить некоторые результаты теории лазера для частного случая, задаваемого неравенством (7.232), и некоторые основные уравнения микроскопической теории.

Предположим пока для большей наглядности, что амплитуда поля b вещественна, и мы можем отождествить ее с переменной ξ, входящей в правую часть соотношения (7.234). Тогда, как показано в теории лазера, $B \equiv \xi$ удовлетворяет уравнению Ланжевена типа (2.1) при условии (2.3). Это уравнение Ланжевена может быть преобразовано в уравнение Фоккера–Планка, стационарное решение которого есть функция (7.234) с параметром

$$\alpha = \frac{\lambda_u(\alpha')}{Q_{\text{полн}}}. \tag{7.235}$$

Здесь $Q_{\text{полн}}$ – полная флуктуирующая сила, действующая на лазерное поле, в то время как $\lambda_u(\alpha')$ – величина, определяемая по формуле

$$\lambda_u = -\frac{\varkappa + \gamma}{2} + \frac{1}{2}\sqrt{(\varkappa + \gamma)^2 + 4g^2\alpha'}, \tag{7.236}$$

при выводе которой мы воспользовались соотношениями (7.226) и (7.229). Разлагая правую часть в ряд по степеням α' и удерживая только главный член разложения, получаем

$$\lambda_u \approx \frac{g^2\alpha'}{\varkappa + \gamma}. \tag{7.237}$$

Сравнивая формулы (7.235) и (7.237), мы получаем соотношение между двумя типами управляющих параметров α и α':

$$\alpha = \frac{g^2}{(\varkappa + \gamma)\, Q_{\text{полн}}}\,\alpha'. \tag{7.238}$$

Используя результат теории лазеров, согласно которому

$$Q_{\text{полн}} \approx \frac{2\varkappa N_{2s}}{D_c}, \tag{7.239}$$

где N_{2s} – число возбужденных атомов, мы можем записать соотношение (7.238) в виде

$$\alpha = \frac{\alpha'}{2N_{2s}} \tag{7.240}$$

(с учетом соотношений (7.229) и (7.232)). Чтобы аппроксимация была хорошей, положим

$$N_{2s} \approx \frac{N}{2}, \tag{7.241}$$

где N – общее число атомов в активной среде лазера.

Соотношения (7.240) и (7.241) позволяют преобразовать формулу (7.231) к следующему виду:

$$\frac{d\lambda_1}{d\alpha}/\lambda_1 = \frac{\varkappa}{\gamma}\frac{N}{D_c} \quad \text{и} \quad \frac{d\lambda_1}{d\alpha} = 0 \quad \text{при} \quad s \geqslant 2. \tag{7.242}$$

Тем самым мы можем записать изменение информации (7.230) как

$$\sum_{s=1}^{N} \left.\frac{dI_s}{d\alpha}\right|_{\alpha=0} = -\frac{\varkappa}{\gamma}\frac{N}{D_c}. \tag{7.243}$$

Отношение D_c/N можно интерпретировать как критическую инверсность на атом d_c. Для характерных значений гелий-неонового лазера мы получаем для комплексной амплитуды лазера

$$\sum_s \left.\frac{dI_s}{d\alpha}\right|_{\alpha=0} \approx 10^6; \qquad \left.\frac{dI_u}{d\alpha}\right|_{\alpha=0} \approx 10^8. \tag{7.244}$$

Глава 8

Прямое определение множителей Лагранжа

8.1. Информационная энтропия систем ниже и выше их критической точки

Из теории неравновесных фазовых переходов известно, что функция распределения параметров порядка упрощается, если система переходит в область, расположенную существенно ниже или выше критической точки. Критические флуктуации системы вблизи критической точки в этом случае уступают место нормальным флуктуациям, описываемым гауссовскими распределениями параметров порядка. Кроме того, ниже критической точки, т. е. в неупорядоченной фазе, параметры порядка ξ_u относительно малы, поэтому в соотношении

$$P(\xi_s|\xi_u) = N\exp\left[-\frac{(\xi_s - g_s(\xi_u))/|\lambda_s|^2}{Q_s}\right] \tag{8.1}$$

мы можем положить $\xi_u = 0$.

Выше порога система (вследствие нарушения симметрии) переходит в состояние $\xi_u = \xi_0 \neq 0$, но при этом флуктуации настолько малы, что в выражении (8.1) мы можем заменить ξ_u на ξ_0. Таким образом, совместная функция распределения $P(\xi_u, \xi_s)$ становится *гауссовской*. В этом разделе мы покажем, каким образом множители Лагранжа можно определить, зная лишь первый и второй моменты. Моменты высшего порядка определяются после этого однозначно. Сравнение вычисленных моментов старшего порядка с их экспериментально измеренными значениями может служить хорошей проверкой применимости гауссовского приближения.

Запишем функцию распределения, множители Лагранжа которой требуется определить, в обычной форме

$$f(\lambda, \mathbf{q}) = \exp\left[-\lambda - \sum_j \lambda_j q_j - \sum_{j,k} \lambda_{jk} q_j q_k\right] \tag{8.2}$$

и сравним ее с функцией распределения

$$\tilde{f}(\mathbf{q}) = (2\pi)^{-n/2}|Q|^{-1/2}\exp\left[-\frac{1}{2}(\mathbf{q}-\mathbf{m})^T Q^{-1}(\mathbf{q}-\mathbf{m})\right], \tag{8.3}$$

где мы воспользовались векторными обозначениями. Верхний индекс

T означает транспонирование; $Q = (Q_{ik})$ – матрица; $|Q|$ – определитель матрицы Q. Нетрудно показать, что функция (8.3) нормирована на единицу в **q**-пространстве. Кроме того, не составляет труда убедиться в том, что

$$m_j = \langle q_j \rangle, \tag{8.4}$$

$$Q_{ij} = \langle (q_i - m_i)(q_j - m_j) \rangle \equiv \langle q_i q_j \rangle - \langle q_i \rangle \langle q_j \rangle. \tag{8.5}$$

Это означает, что, коль скоро моменты $\langle q_i \rangle$ и $\langle q_i q_j \rangle$ известны (или заданы), мы можем определить **m** и Q, входящие в правую часть соотношения (8.3).

Следовательно, чтобы определить множители Лагранжа в функции распределения (8.2), необходимо установить соотношение, с одной стороны, между **m** и Q, а с другой стороны, между λ, λ_j и λ_{jk}. Отдельные этапы вычислений тривиальны, и окончательный результат оказывается следующим:

$$\exp[-\lambda] = (2\pi^{-n/2})|Q|^{-1/2}\exp\left[-\frac{1}{2}\mathbf{m}^T Q^{-1}\mathbf{m}\right], \tag{8.6}$$

$$\lambda_{jk} = \frac{1}{2}[Q^{-1}]_{jk}, \tag{8.7}$$

$$\lambda_j = -\frac{1}{2}\sum_k m_k\{[Q^{-1}]_{kj} + [Q^{-1}]_{jk}\}. \tag{8.8}$$

Мы можем воспользоваться теперь формулой (3.47), которую запишем в виде

$$i = \lambda + \sum_j \lambda_j \langle q_j \rangle + \sum_{j,k} \lambda_{jk} \langle q_j q_k \rangle. \tag{8.9}$$

Подставляя выражения (8.6)–(8.8) в формулу (8.9) и учитывая соотношения (8.4) и (8.5), получаем

$$i = \frac{n}{2}\ln 2\pi + \frac{1}{2}\ln|Q| + \frac{1}{2}\sum_{j,k}[Q^{-1}]_{jk}\underbrace{(\langle q_j q_k \rangle - \langle q_j \rangle \langle q_k \rangle)}_{Q_{kj}}. \tag{8.10}$$

Но матричное соотношение

$$\sum_{j,k}[Q^{-1}]_{jk} Q_{kj} = \sum_j (Q^{-1}Q)_{jj} = \sum_j 1 = n \tag{8.11}$$

позволяет нам получить окончательный результат – информацию, соответствующую функции распределения (8.2) с ограничениями (8.4) и (8.5):

$$i = \frac{n}{2}\ln 2\pi + \frac{1}{2}\ln|Q| + \frac{n}{2}. \quad (8.12)$$

Совершенно очевидно, что первые моменты $\langle q_i \rangle$ выпадают.

Как показывает сравнение с формулой (7.61), соотношение (8.12) может быть интерпретировано как сумма по n информационных энтропий независимых гауссовских распределений с ширинами $1/\alpha_k$, где

$$\ln|Q| = -\sum_k \ln 2\alpha_k. \quad (8.13)$$

Упражнение. Проверьте нормировку функции $\tilde{f}(\mathbf{q})$, задаваемую соотношением (8.3) и соотношениями (8.4) и (8.5).

Указание: диагонализуйте матрицу Q с помощью преобразования $\mathbf{q} = A\xi$. Проверьте соотношения (8.6)–(8.8).

Указание: раскройте скобки в выражении (8.3) и сравните член за членом выражения, стоящие под знаком экспоненты в формулах (8.2) и (8.3).

8.2. Прямое определение множителей Лагранжа ниже критической точки, в критической точке и выше критической точки

В этом разделе мы хотим показать, как можно, не прибегая к интегрированию, определить множители Лагранжа, которые необходимы при использовании принципа максимума информационной энтропии, не прибегая к вычислению интегралов. За прямой подход приходится платить следующую цену: необходимо предположить, что мы имеем дело со специфическим переходом, например с переходом мягкой моды, о котором шла речь в разд. 7.4–7.9. Разобьем процедуру на два этапа. Начнем с ограничений вида

$$f_j = \langle q_j \rangle, \qquad f_{jj'} = \langle q_j q_{j'} \rangle, \ \ldots, \quad (8.14)$$

содержащих моменты до четвертого порядка, преобразуем переменные q_j в новые переменные и определим соответствующие преобразования. Затем учтем, что мы рассматриваем все для случая неустойчивости мягкой моды, и покажем, каким образом могут быть прямо, т.е. без интегрирования, определены множители Лагранжа. Одновременно мы получим условия самосогласованности для ограничений (8.14). Эти условия позволяют проверить, претерпевает ли рассматриваемая система мягкомодовый переход в соответствии с нашей первоначальной гипотезой.

Чтобы упростить вычисления, введем следующие обозначения. Взяв, например, $f_{jj'}$ в форме (8.14), образуем матрицу и условимся рассматривать ее как прямое произведение вектора $\mathbf{q}$ на себя

$$\langle \mathbf{q} \circ \mathbf{q} \rangle = (\langle q_j q_{j'} \rangle). \tag{8.15}$$

В случае трех индексов j, j', j'' мы можем образовать тройное произведение вектора $\mathbf{q}$ на себя

$$\langle \mathbf{q} \circ \mathbf{q} \circ \mathbf{q} \rangle = (\langle q_j q_{j'} q_{j''} \rangle), \tag{8.16}$$

где в правой части необходимо перебрать все комбинации индексов j, j'' и j'''. Аналогичным образом мы можем поступить и в случае моментов четвертого порядка. Обратимся теперь к микроскопическому подходу, в котором вектор состояния $\mathbf{q}$ был представлен в виде суперпозиции векторов

$$\mathbf{q} = \mathbf{q}_0 + \sum \xi_k \mathbf{v}_k, \tag{8.17}$$

где $\mathbf{q}_0$ – постоянный вектор состояния, который мы в дальнейшем будем считать равным нулю:

$$\mathbf{q}_0 = \mathbf{0}. \tag{8.18}$$

Нетрудно, однако, рассмотреть и такой случай, когда $\mathbf{q}_0 \neq \mathbf{0}$. Коэффициенты ξ_k – случайные величины, $\mathbf{v}_k$ – постоянные, пока еще не определенные векторы. Введем сопряженный вектор

$$\bar{\mathbf{v}}_k = (v_{k1}^*, v_{k2}^*, \dots, v_{kn}^*) \tag{8.19}$$

и предположим, что выполняется соотношение ортогональности

$$\bar{\mathbf{v}}_k \bar{\mathbf{v}}_{k'} = \delta_{kk'}. \tag{8.20}$$

Усредняя разложение (8.17), получаем

$$\langle \mathbf{q} \rangle = \sum_k \langle \xi_k \rangle \mathbf{v}_k. \tag{8.21}$$

Умножая обе части соотношения (8.21) на сопряженный вектор $\bar{\mathbf{v}}_k$, приходим к соотношению

$$\bar{\mathbf{v}}_k \langle \mathbf{q} \rangle = \langle \xi_k \rangle \tag{8.22}$$

(вследствие условия ортогональности (8.20)). Вводя для краткости обозначение

$$C_k = \mathbf{v}_k \langle \mathbf{q} \rangle, \tag{8.23}$$

мы можем записать соотношение (8.22) в следующем виде:

$$\langle \xi_k \rangle = C_k. \tag{8.24}$$

Умножая подходящим образом равенство (8.15) на сопряженные векторы $\bar{\mathbf{v}}$ и используя соотношение ортогональности (8.20), получаем

$$\langle \xi_k \xi_{k'} \rangle = \langle \bar{\mathbf{v}}_k \mathbf{q} \cdot \mathbf{v}_{k'} \mathbf{q} \rangle \equiv C_{kk'}, \tag{8.25}$$

где последнее равенство служит определением величины $C_{kk'}$. Средняя часть равенств (8.25) определяется в более подробной записи соотношением

$$\langle \bar{\mathbf{v}}_k \mathbf{q} \cdot \bar{\mathbf{v}}_{k'} \mathbf{q} \rangle = \sum_l \mathbf{v}^*_{kl} \mathbf{q}_l \cdot \sum_{l'} \mathbf{v}^*_{k'\,l'} \mathbf{q}_{l'}. \tag{8.26}$$

Точно так же мы находим, что для третьего и четвертого моментов справедливы соотношения

$$\langle \xi_k \xi_{k'} \xi_{k''} \rangle = \langle \bar{\mathbf{v}}_k \mathbf{q} \cdot \bar{\mathbf{v}}_{k'} \mathbf{q} \cdot \bar{\mathbf{v}}_{k''} \mathbf{q} \rangle = C_{k\,k'\,k''}, \tag{8.27}$$

$$\langle \xi_k \xi_{k'} \xi_{k''} \xi_{k'''} \rangle = \langle \prod_{\nu=1}^{4} \bar{\mathbf{v}}_k \mathbf{q} \rangle = C_{k\,k'\,k''\,k'''}. \tag{8.28}$$

До сих пор мы занимались преобразованиями, переводящими старые моменты, выраженные через $\mathbf{q}$, в новые моменты, выраженные через ξ_k. В дальнейшем мы будем предполагать, что старые моменты (8.14) известны из эксперимента, и определим новые моменты

$$\langle \xi_k \xi_{k'} \ldots \rangle \tag{8.29}$$

и векторы $\bar{\mathbf{v}}_k$. Для этого мы предположим теперь в явном виде, что рассматриваем и описываем моментами (8.14) такое явление, как неустойчивость мягкой моды. Для дальнейшего нам необходимо напомнить читателю о подходе на микроскопическом уровне. Ниже порога система очень хорошо аппроксимируется гауссовским распределением по всем переменным системы. Это позволяет ограничиться первым и вторым моментом. Чтобы найти подходящие векторы мод $\bar{\mathbf{v}}_k$, можно диагонализовать матрицу

$$\langle \mathbf{q} \circ \mathbf{q} \rangle. \tag{8.30}$$

Диагонализуя ее, мы получаем вещественные собственные значения, соответствующие вторым моментам мод. Как нам уже известно, в случае неустойчивости мягкой моды одна из них испытывает при приближении к критической точке критические флуктуации. Следовательно, можно ожидать, что одно собственное значение станет очень большим, в то время как все остальные останутся малыми. Если какое-то собственное значение становится большим, то соответствующая ему мода неустойчива, а другие собственные значения должны отвечать устойчивым модам. Так мы можем отличать мягкую моду от устойчивых мод.

Следует подчеркнуть, однако, что в отличие от математической линеаризации, при которой собственное значение мягкой моды обращается в бесконечность, в рассматриваемом случае собственное значение остается ограниченным, поскольку матрица (8.30) построена по экспериментальным данным, а в реальном эксперименте нелиней-

ности не дают собственным значениям возрастать до бесконечности.

Введем теперь различие между неустойчивой и устойчивой модами:

$$\mathbf{v}_j \to \mathbf{v}_u,\ \mathbf{v}_s. \tag{8.31}$$

Рассмотрим типичный переход мягкой моды, для которого из соображений симметрии можно записать равенство

$$\langle \xi_u \rangle = 0. \tag{8.32}$$

В дальнейшем мы будем дополнительно предполагать, что

$$\langle \xi_u \xi_s \rangle = 0. \tag{8.33}$$

Если говорить точнее, то мы будем предполагать, что совместная функция распределения, по которой производится усреднение, имеет вид

$$P(\xi_u,\ \xi_s) = f(\xi_u) \prod_s P_s(\xi_s \mid \xi_u). \tag{8.34}$$

Формулу (8.34) мы можем заимствовать либо из микроскопического подхода, либо из того макроскопического подхода, который был изложен в гл. 6. Здесь f – функция распределения параметра порядка, а P_s – условная вероятность для амплитуды ξ_s подчиненной моды, задаваемая соотношением

$$P = N_s \exp\left\{ -\frac{[\xi_s - g_s(\xi_u)]^2}{Q_s} \right\}. \tag{8.35}$$

В дальнейшем мы рассмотрим один частный, но весьма частый случай, когда g_s – четная функция параметра порядка. Преобразование

$$\xi_s = \bar{\xi}_s + g_s(\xi_u) \tag{8.36}$$

позволяет легко вычислять средние значения, например

$$\langle \xi_s \rangle = \langle \bar{\xi}_s + g_s(\xi_u) \rangle = \langle \bar{\xi}_s \rangle + \langle g_s(\xi_u) \rangle. \tag{8.37}$$

С учетом соотношения (8.35) получаем условие

$$\langle \bar{\xi}_s \rangle = 0. \tag{8.38}$$

В дальнейшем мы будем для краткости записывать вместо переменных ξ_s и ξ_u только их индексы s и u:

$$\xi_s \to s, \qquad \xi_u \to u. \tag{8.39}$$

В этих новых обозначениях соотношение (8.37) имеет следующий вид:

$$\langle s \rangle = \langle \bar{s} + g_s \rangle = \langle \bar{s} \rangle + \langle g_s \rangle. \tag{8.40}$$

Исследуем теперь выражения (8.24), (8.25), (8.27) и (8.28), имея в виду, что функция распределения выбрана нами вполне определенного вида, а именно функция, задаваемая выражениями (8.34) и (8.35). Из формул (8.32), (8.23) и (8.24) мы без труда получаем соотношение

$$\bar{\mathbf{v}}_u(\mathbf{q}) = 0 \tag{8.41}$$

(это новое условие на среднее значение вектора $\mathbf{q}$), а также соотношение

$$\langle s \rangle = \langle g_s(u) \rangle = C_s. \tag{8.42}$$

Как мы увидим из дальнейшего, оно позволяет при соответствующих предположениях определить функцию g_s. Из соотношения (8.25), полагая $k = u$, $k' = u$, находим

$$\langle u^2 \rangle = C_{uu}. \tag{8.43}$$

Полагая $k = u$, $k' = s$ и используя выражения (8.34) и (8.35), получаем

$$\langle us \rangle = 0 = C_{us}, \tag{8.44}$$

что есть условие на C_{us}. Наконец, полагая $k = s$, $k' = s'$, приходим к соотношению

$$\langle ss' \rangle = \langle \bar{s}^2 \rangle \delta_{ss'} + \langle g_s g_{s'} \rangle = C_{ss'}. \tag{8.45}$$

Поскольку

$$\langle \bar{s}^2 \rangle = Q_s, \tag{8.46}$$

соотношение (8.45) можно записать в виде

$$Q_s \delta_{ss'} + \langle g_s g_{s'} \rangle = C_{ss'}. \tag{8.47}$$

Как мы увидим из дальнейшего, этим соотношением можно воспользоваться для определения величины Q_s. Выясним теперь, какие следствия можно извлечь из соотношения

$$\langle \xi_k \xi_{k'} \xi_{k''} \rangle = C_{k k' k''}. \tag{8.48}$$

1) Пусть

$$k = u, \quad k' = s', \quad k'' = s'' \tag{8.49}$$

(возможна также любая перестановка величин, стоящих в правых частях равенств (8.49)). Тогда

$$C_{u s' s''} = 0. \tag{8.50}$$

Это соотношение следует из свойств левой части соотношения (8.48). Оно также принадлежит к числу соотношений, которым должны удовлетворять экспериментально найденные моменты.

2) Пусть

$$k = u, \quad k' = u, \quad k'' = s'' \tag{8.51}$$

(возможна также любая перестановка величин, стоящих в правых частях равенств (8.51)). Тогда

$$\langle \xi_u^2 g_{s''}(\xi_u) \rangle = C_{u\,u\,s''}. \tag{8.52}$$

3) Пусть

$$k = u, \quad k' = u, \quad k'' = u. \tag{8.53}$$

Тогда

$$C_{u\,u'\,u''} = 0. \tag{8.54}$$

Равенство (8.54) также принадлежит к числу условий, которым должны удовлетворять экспериментально найденные моменты.

4) Пусть

$$k = s, \quad k' = s', \quad k'' = s''. \tag{8.55}$$

Воспользуемся теперь разложением (8.36) и запишем его в тех же обозначениях, что и соотношение (8.40). Тогда вместо соотношения (8.48) мы получим новое соотношение

$$\langle (\bar{s} + g_s)(\bar{s}' + g_{s'})(\bar{s}'' + g_{s''}) \rangle. \tag{8.56}$$

Раскрывая круглые скобки и приводя подобные члены, приходим к соотношению

$$\langle \bar{s}\bar{s}'\bar{s}'' \rangle = 0. \tag{8.57}$$

Кроме того, мы получаем соотношения

$$\langle \bar{s}\bar{s}' g_{s''} \rangle = \delta_{ss'} \langle \bar{s}^2 \rangle \langle g_{s''} \rangle, \tag{8.58}$$

$$\langle \bar{s} g_{s'} g_{s''} \rangle = 0 \tag{8.59}$$

и, наконец, находим выражение

$$\langle g_s g_{s'} g_{s''} \rangle, \tag{8.60}$$

которое пока не допускает дальнейшего упрощения. Собирая члены (8.57)–(8.60) (вместе с величинами, получающимися из них при соответствующих перестановках), приходим к соотношению

$$Q_s \delta_{ss'} \langle g_{s''} \rangle + Q_s \delta_{ss''} \langle g_{s'} \rangle + Q_{s'} \delta_{s's''} \langle g_s \rangle + \langle g_s g_{s'} g_{s''} \rangle = C_{s\,s'\,s''}, \tag{8.61}$$

которое следует из соотношения (8.48). Для полноты приведем также соотношения, которые следуют из равенства

$$\langle \xi_k \xi_{k'} \xi_{k''} \xi_{k'''} \rangle = Q_{k\,k'\,k''\,k'''}. \tag{8.62}$$

Нетерпеливый читатель может опустить все промежуточные вычисления и перейти сразу к окончательному результату – формуле (8.83).

Придавая индексам k, k', k'', k''' различные значения, мы выделяем несколько случаев.

1. Если индексы принимают значения

$$u,\ u',\ u'',\ u''', \tag{8.63}$$

то

$$\langle u^4 \rangle = C_{uuuu}. \tag{8.64}$$

2. Если индексы принимают значения

$$u,\ u,\ u,\ s'', \tag{8.65}$$

то

$$C_{uuus''} = 0. \tag{8.66}$$

3. Если индексы принимают значения

$$u,\ u,\ s'',\ s''', \tag{8.67}$$

то, производя разложение (8.36), которое дает

$$\langle u^2 (\bar{s}'' + g_{s''})(\bar{s}''' + g_{s'''}) \rangle, \tag{8.68}$$

и приводя подобные члены, получаем

$$\langle u^2 \bar{s}'' \bar{s}''' \rangle = \langle u^2 \rangle Q_{s''} \delta_{s'' s'''}, \tag{8.69}$$

$$\langle u^2 \bar{s}'' g_{s'''} \rangle = 0, \tag{8.70}$$

$$\langle u^2 g_{s''} g_{s'''} \rangle \neq 0. \tag{8.71}$$

4. Если индексы принимают значения

$$u,\ s',\ s'',\ s''', \tag{8.72}$$

то по соображениям симметрии, как упоминалось выше,

$$C_{u s' s'' s'''} = 0. \tag{8.73}$$

5. Если индексы принимают значения

$$s,\ s',\ s'',\ s''', \tag{8.74}$$

то, производя соответствующее разложение, мы получаем выражение

$$\langle (\bar{s} + g_s)(\bar{s}' + g_{s'})(\bar{s}'' + g_{s''})(\bar{s}''' + g_{s'''}) \rangle \tag{8.75}$$

и можем рассматривать его отдельные члены. Чтобы вычислить член

$$\langle \bar{s} \bar{s}' \bar{s}'' \bar{s}''' \rangle, \tag{8.76}$$

необходимо учитывать такие комбинации индексов s, которые порождают ненулевые средние значения. Этим свойством обладают те комбинации, в которых две пары индексов совпадают. Иначе говоря, должны выполняться соотношения

$$s = s', \quad s'' = s''';$$
$$s = s'', \quad s' = s''';$$
$$s = s''', \quad s' = s''. \tag{8.77}$$

Таким образом, вместо члена (8.76) мы получаем следующее выражение:

$$\langle \bar{s}\bar{s}'\bar{s}''\bar{s}''' \rangle = \langle \bar{s}^2 \rangle \langle \bar{s}''^2 \rangle \delta_{ss'} \delta_{s''s'''} + \langle \bar{s}^2 \rangle \langle \bar{s}'^2 \rangle \delta_{ss''} \delta_{s's'''} + + \langle \bar{s}^2 \rangle \langle \bar{s}'^2 \rangle \delta_{ss'''} \delta_{s's''}. \tag{8.78}$$

Следующий член, который мы должны рассмотреть, типа

$$\langle \bar{s}\bar{s}'\bar{s}'' g_{s'''} \rangle = 0 \tag{8.79}$$

и членов, получающихся при перестановках, в которых множитель g заменяет один из s-членов. Следующий член имеет вид

$$\langle \bar{s}\bar{s}' g_{s''} g_{s'''} \rangle = \langle \bar{s}^2 \rangle \delta_{ss'} \langle g_{s''} g_{s'''} \rangle = Q_s \delta_{ss'} \langle g_{s''} g_{s'''} \rangle, \tag{8.80}$$

где для получения ненулевого среднего необходимо положить $\bar{s} = \bar{s}'$. Переставляя индексы, находим другие члены того же типа. Для членов вида

$$\langle \bar{s} g_{s'} g_{s''} g_{s'''} \rangle = 0 \tag{8.81}$$

результат оказывается таким, как говорилось выше. Наконец, существует и такой член, как

$$\langle g_s g_{s'} g_{s''} g_{s'''} \rangle. \tag{8.82}$$

Собирая все члены от (8.64) до (8.82), записываем окончательный результат в виде соотношения

$$Q_s Q_{s''} \delta_{ss'} \delta_{s''s'''} + Q_s Q_{s'} \delta_{ss''} \delta_{s's'''} + Q_s Q_{s'} \delta_{ss'''} \delta_{s's''} + + Q_s [\delta_{ss'} \langle g_{s''} g_{s'''} \rangle + \delta_{ss''} \langle g_{s'} g_{s'''} \rangle + \delta_{ss'''} \langle g_{s'} g_{s''} \rangle] + Q_{s'} [\delta_{s's''} \langle g_s g_{s'''} \rangle + + \delta_{s's'''} \langle g_s g_{s''} \rangle] + Q_{s''} \delta_{s''s'''} \langle g_s g_{s'} \rangle + \langle g_s g_{s'} g_{s''} g_{s'''} \rangle = C_{ss's''s'''}, \tag{8.83}$$

которое следует из соотношения (8.62) для случая, когда все индексы s-типа. Рассмотрим теперь наши результаты, чтобы зафиксировать неизвестные результаты, входящие в качестве множителей Лагранжа в функцию распределения (8.34) и (8.35). Свой анализ мы ограничим областью, расположенной не слишком высоко над порогом. В этой области величину g_s допустимо аппроксимировать главным членом, а именно

$$g_s(u) \approx a_s u^2. \tag{8.84}$$

Примем гипотезу о форме функции распределения параметров по-

рядка, а именно будем считать, что функция распределения определяется выражением

$$f(u) = \mathcal{N} \exp(\alpha u^2 + \beta u^4) \tag{8.85}$$

в соответствии с приведенными выше результатами из микроскопической и макроскопической теории. Таким образом, требуется определить следующие константы: α, β, Q_1, a_s. Постоянные α, β и нормирующий множитель $\mathcal{N}$ определяются соотношениями (8.43) и (8.64), т. е. соотношениями

$$\langle u^2 \rangle = C_{uu}, \tag{8.86}$$

$$\langle u^4 \rangle = C_{uuuu}, \tag{8.87}$$

а постоянная a_s может быть определена соотношением (8.42), т. е.

$$a_s \langle u^2 \rangle = C_s. \tag{8.88}$$

Наконец, постоянная Q_s определяется соотношением (8.47) при $s = s'$ и соотношением

$$\langle g_s^2 \rangle = a_s^2 \langle u^4 \rangle, \tag{8.89}$$

которое следует из формулы (8.84). Таким образом, постоянная Q_s может быть вычислена по формуле

$$Q_s = C_{ss} - \frac{C_s^2 \langle u^2 \rangle^2}{\langle u^4 \rangle}. \tag{8.90}$$

В результате мы получаем возможность определить все неизвестные постоянные. Нетрудно убедиться, что тем самым мы находим множители Лагранжа, которые входят в выражение (3.38), или, что эквивалентно, в выбранную нами конкретную форму функции распределения (8.34), (8.35), (8.84) и (8.85). Следует подчеркнуть, что система уравнений, которым удовлетворяют интересующие нас константы, переопределена, т. е. уравнений больше, чем неизвестных. Поэтому остальные соотношения, выведенные выше и содержащие экспериментально измеренные величины C, являются условиями самосогласованности для существования неравновесного фазового перехода с одной мягкой модой.

Полученные результаты нетрудно обобщить на более сложные случаи, например на случай, когда соотношение (8.84) содержит нечетные и/или более высокие степени параметра ξ_u, или на случай нескольких параметров порядка.

Глава 9
Непредвзятое моделирование стохастических процессов: как легко и просто получать интегралы по траекториям, уравнения Фоккера–Планка и уравнения Ланжевена–Ито

9.1. Одномерный вектор состояния

В предыдущих главах мы показали, как принцип максимума информационной энтропии позволяет нам получать функции распределения неравновесных систем с помощью подходящего выбора ограничений. При этом наше внимание было направлено на стационарные функции распределения. В этой главе мы намереваемся исследовать нестационарные функции распределения. Исходным пунктом для наших построений служит так называемый принцип максимального калибра, сформулированный Джейнсом. Предложенный им же принцип максимума энтропии Джейнс обобщил на случай, когда максимизации по некоторой области пространства-времени подлежит функционал типа энтропии. Свой принцип Джейнс использовал для вывода важных формул, применимых к термодинамике необратимых процессов для исследования релаксационных процессов.

Мы хотим показать, что адекватный рассматриваемой задаче выбор ограничений позволяет рассматривать любую физическую (или какую-нибудь другую) систему как вблизи, так и вдали от равновесия. Предлагаемая нами процедура допускает особенно изящную формулировку, если лежащий в основе рассматриваемой задачи стохастический процесс является непрерывным марковским процессом. В этом случае мы очень быстро приходим к интегралам, их траекториям и уравнению Фоккера–Планка, в котором коэффициент дрейфа и коэффициент диффузии могут быть однозначно определены. Таким образом, наш метод можно рассматривать как способ, позволяющий находить детерминистические и флуктуирующие силы, действующие на систему, по экспериментальным данным.

Рассмотрим последовательность моментов времени $t_0, t_1, \ldots, t_N$, в которые мы производим измерения над системой. Пусть q_i – значение вектора состояния, измеренное в момент времени t_i. Мы хотим получить несмещенную оценку совместной функции распределения

$$P_N = P(\mathbf{q}_N, t_N; \mathbf{q}_{N-1}, t_{N-1}; \ldots; \mathbf{q}_0, t_0). \tag{9.1}$$

Для этого мы найдем максимум информации i, или, если следовать терминологии Джейнса, калибра

$$i = -\int \mathcal{D}q P \ln P, \tag{9.2}$$

по P при заданных ограничениях. Величина $\mathcal{D}q$ означает здесь

элемент объема, по которому интегрируем $P \ln P$ по пространству, натянутому на все векторы $\mathbf{q}_j$.

Как хорошо известно, в задачах этого типа решающую роль играет адекватный выбор ограничений. Этот выбор сильно облегчается, если предположить, что лежащий в основе задачи стохастический процесс, который приводит к функции распределения (9.1), является марковским. В этом случае функция распределения (9.1) допускает разложение

$$P_N = \prod_{l=0}^{N-1} \mathrm{P}(\mathbf{q}_{l+1}, t_{l+1} \,|\, \mathbf{q}_l, t_l) \cdot P_0(\mathbf{q}_0, t_0). \tag{9.3}$$

Чтобы упростить обозначения, условимся опускать моменты времени t_l, тогда отдельные сомножители в формуле (9.3) примут вид

$$P(\mathbf{q}_{i+1} \,|\, \mathbf{q}_i). \tag{9.4}$$

Выражение (9.4) определяет условную вероятность. В первой части этой главы мы будем предполагать, что стохастический процесс стационарен.

Чтобы пояснить суть нашего метода, рассмотрим вместо вектора состояния одну переменную q. Введем ограничения

$$f_1 = \langle q_{i+1} \rangle_{q_i}, \tag{9.5}$$

$$f_2 = \langle q_{i+1}^2 \rangle_{q_i}, \tag{9.6}$$

определяемые соотношением

$$f_k = \int P(q_{i+1} \,|\, q_i)\, q_{i+1}^k \, dq_{i+1}. \tag{9.7}$$

Запись (9.7) означает следующее: мы рассматриваем первый и второй моменты величины q в момент времени t_{i+1} при условии, что значение q_i было измерено в момент времени t_i. Применяя к условной вероятности (9.4) принцип максимума информационной энтропии с ограничениями (9.5) и (9.6), получаем

$$P(q_{i+1} \,|\, q_i) = \exp(\lambda + \lambda_1 q_{i+1} + \lambda_2 q_{i+1}^2). \tag{9.8}$$

Заметим, что f_1 и f_2 зависят от q_i, поэтому λ, λ_1 и λ_2 по крайней мере в общем случае так же зависят от q_i. Так как мы имеем в виду предельный переход $\tau \to 0$, условимся в дальнейшем вместо $q(i+1)$ писать $q(i+\tau)$. Наложим на условную вероятность (9.8) обычные требования, а именно условия нормировки на единицу

$$\int P(q_{i+\tau} \,|\, q_i)\, dq_{i+\tau} = 1. \tag{9.9}$$

Кроме того, потребуем, чтобы при $\tau \to 0$

$$P(q_{i+\tau} \,|\, q_i) \to \delta(q_{i+\tau} - q_i). \tag{9.10}$$

Дельта-функцию, стоящую в правой части соотношения (9.10), следует понимать особым образом, а именно так, что условная вероят-

ность P ведет себя как δ-функция при интегрировании по $q_{i+\tau}$.

Чтобы более полно использовать условия (9.9) и (9.10), запишем правую часть формулы (9.8) как

$$\exp\left[-|\lambda_2|\left(q_{i+\tau}-\frac{\lambda_1}{2|\lambda_2|}\right)^2-\frac{\lambda_1^2}{4|\lambda_2|}+\lambda\right]. \tag{9.11}$$

Так как

$$\int \exp(-|\lambda_2|q^2)\,dq=\sqrt{\pi/|\lambda_2|}, \tag{9.12}$$

условие нормировки (9.9) можно записать в виде

$$\exp\left(\lambda-\frac{\lambda_1^2}{4|\lambda_2|}\right)=\sqrt{\frac{|\lambda_2|}{\pi}}, \tag{9.13}$$

вследствие чего условная вероятность P принимает вид

$$P(q_{i+\tau}|q_i)=\sqrt{|\lambda_2(q_i)|/\pi}\exp\left[-|\lambda_2|\left(q_{i+\tau}-\frac{\lambda_1}{2|\lambda_2|}\right)^2\right]. \tag{9.14}$$

Чтобы при $\tau\to 0$ вероятность (9.14) стала δ-функцией, величина $|\lambda_2|$ должна вести себя как

$$|\lambda_2|=\frac{G}{\tau}. \tag{9.15}$$

(Другая зависимость от τ, при которой величина $|\lambda_2|$ обращается в нуль при $\tau\to 0$, означает просто другую зависимость от времени, при которой сама функция, задающая закон, по которому происходит скейлинг, зависит от времени.)

Из условия (9.10) следует, что при $\tau\to 0$

$$q_{i+\tau}-\frac{\lambda_1}{2|\lambda_2|}\to q_{i+\tau}-q_i, \tag{9.16}$$

или

$$\frac{\lambda_1}{2|\lambda_2|}\to q_i. \tag{9.17}$$

Формула предельного перехода (9.17) при малых, но отличных от нуля временах τ допускает очевидное обобщение:

$$\frac{\lambda_1}{2|\lambda_2|}=q_i+\tau K(q_i)+\tau^2 H(q_i)+\ldots. \tag{9.18}$$

Таким образом, мы можем утверждать, что величина $\lambda_1/2|\lambda_2|$, стоящая в левой части соотношения (9.18), при малых временах τ должна вести себя как аналитическая функция параметра τ. Имея соотношения (9.14) и (9.18), мы получаем главный результат этого

раздела, а именно выражение для условной вероятности в пределах малых времен:

$$P(q_{i+\tau} \mid q_i) = \sqrt{G/(\tau\pi)} \exp\left\{ -\frac{G}{\tau} [q_{i+\tau} - q_i - \tau K(q_i)]^2 \right\}. \tag{9.19}$$

Заметим, что G может также зависеть от q_i.

9.2. Обобщение на случай многомерного вектора состояния

В этом разделе мы хотим получить несмещенную оценку вероятности (9.3) в предположении, что основной процесс является стационарным непрерывным марковским процессом. Обозначим компоненты векторов состояния в моменты времени $i + \tau$ и i через

$$q_l(i + \tau) \quad \text{и} \quad q_k(i). \tag{9.20}$$

В качестве обобщения первого и второго моментов одной переменной (9.5) и (9.6) введем следующие первый и второй условные моменты:

$$f_{1,l} = \langle q_l(i + \tau) \rangle_{\mathbf{q}(i)}, \tag{9.21}$$

$$f_{2,l,k} = \langle q_l(i + \tau)\, q_k(i + \tau) \rangle_{\mathbf{q}(i)}. \tag{9.22}$$

Применяя к условной вероятности принцип максимума информационной энтропии, получаем

$$P(\mathbf{q}(i + \tau) \mid \mathbf{q}(i)) = \exp\left[\lambda + \sum_l \lambda_l q_l(i + \tau) + \sum_{k,l} \lambda_{kl} q_k(i + \tau)\, q_l(i + \tau)\right], \tag{9.23}$$

где множители Лагранжа также могут зависеть от $\mathbf{q}(i)$. В дальнейшем мы будем предполагать, что рассматриваемая нами задача невырожденная, т. е. что

$$\operatorname{Det} \lambda_{kl} \neq 0. \tag{9.24}$$

Несложные алгебраические преобразования показывают, что выражение (9.23) можно привести к виду

$$P = \exp\left\{\tilde{\lambda} + \sum_{k,l} [q_k(i + \tau) - h_k]\, \lambda_{kl} [q_l(i + \tau) - h_l]\right\}, \tag{9.25}$$

где h_l определяется соотношением

$$h_l = \sum_m A_{lm} \lambda_m [\mathbf{q}(i)]. \tag{9.26}$$

Обозначим матрицу с элементами λ_{kl} через Δ. Матрица коэффициентов A в правой части соотношения (9.26) удовлетворяет соотношению

$$A = \frac{1}{2}\Delta^{-1}. \tag{9.27}$$

Заметим, что матрица Δ симметрична. Кроме того, входящая в соотношение (9.25) величина λ определяется соотношением

$$\tilde{\lambda} = \lambda - \sum_{k,\,l} h_k \lambda_{kl} h_l. \tag{9.28}$$

Обобщая условие (9.10), мы требуем, чтобы при $\tau \to 0$ вероятность P вела себя как δ-функция, если интегрирование производится по $\mathbf{q}(i + \tau)$:

$$P \to \delta\,[\mathbf{q}(i + \tau) - \mathbf{q}(i)]. \tag{9.29}$$

Сингулярное поведение δ-функции мы достигаем тем, что

$$\lambda_{kl} \quad \frac{1}{\tau} G_{kl}\,[\mathbf{q}(i)]. \tag{9.30}$$

Обобщая соотношение (9.18), разложим h_k в степенной ряд по τ и удержим только первый член разложения:

$$h_k = q_k(i) + \tau K_k\,[\mathbf{q}(i)] + \ldots. \tag{9.31}$$

Объединяя соотношения (9.25), (9.28), (9.30) и (9.31), мы можем записать условную вероятность в виде

$$P = \exp\,[\tilde{\lambda}(\mathbf{q}_i)]\exp\left\{-\frac{1}{\tau}\sum_{k,\,l}[q_k(i + \tau) - q_k(i) + \tau K_k] \times \right.$$
$$\left. \times\; G_{kl}\,[q_l(i + \tau) - q_l(i) - \tau K_l]\right\}. \tag{9.32}$$

Обозначая для краткости первый множитель в правой части соотношения (9.32) через $\mathcal{N}(\mathbf{q}_i)$, мы можем записать это соотношение более компактно как

$$P = \mathcal{N}(\mathbf{q}_i)\exp(\mathcal{L}_i). \tag{9.33}$$

Подставляя выражение (9.32) в разложение (9.3), где P_0 – заданная функция распределения, относительно которой мы в случае необходимости можем принимать те или иные гипотезы, задавая некоторое количество моментов вектора $\mathbf{q}_0$, приходим к окончательному результату, а именно к совместной функции распределения (9.1).

В практических приложениях интерес представляет главным образом условная вероятность – функция распределения для $\mathbf{q}$ в момент времени T при условии, что вектор состояния $\mathbf{q}$ был известен в момент времени $T = 0$. Интересующая нас вероятность получается при интегрировании по всем $\mathbf{q}$ в промежуточные моменты времени. Полагая

$$N\tau = T \tag{9.34}$$

и идентифицируя величину $\mathscr{L}_i$ с помощью соотношений (9.33) и (9.32), получаем требуемый результат в виде

$$P(\mathbf{q}(T)\,|\,\mathbf{q}(0)) = \int \mathscr{D}q \exp\left(\int_{i=0}^{N-1} \mathscr{L}_i\right), \tag{9.35}$$

где $\mathscr{N}(\mathbf{q}_i)$ содержится в $\mathscr{D}q$.

В пределе при $\tau \to 0$ сумма по i переходит в интеграл, и формула (9.35) есть не что иное, как представление условной вероятности P в виде интеграла по траекториям. Из сказанного в разд. 2.5 о представлении решения уравнения Фоккера–Планка в виде интеграла по траекториям ясно, что выражение (9.35) вместе с формулами (9.33) и (9.32) есть решение уравнения Фоккера–Планка с дрейфовым коэффициентом K_l и матрицей диффузии, задаваемой обратной матрицей с элементами G_{kl}.

Для наших целей достаточно знать, что выражение (9.35) есть решение уравнения Фоккера–Планка с дрейфовым коэффициентом и коэффициентом диффузии, о которых мы только что говорили (см. разд. 9.4).

В заключение этого раздела упомянем о том, что наш подход легко может быть обобщен на случай нестационарного непрерывного марковского процесса. В этом случае, вообще говоря, следует произвести подстановки

$$\mathbf{K}[\mathbf{q}(i)] \to \mathbf{K}[\mathbf{q}(i);\quad t_i], \tag{9.36}$$

$$G_{kl}[\mathbf{q}(i)] \to G_{kl}[\mathbf{q}(i);\quad t_i]. \tag{9.37}$$

9.3. Корреляционные функции как ограничения

В предыдущих разделах мы использовали условные моменты, т. е. моменты переменных $q_{i+\tau}$ при условии, что переменная q_i измерена. В ряде практических приложений нам могут быть известны не q_i, а корреляционные функции типа

$$\langle q_{i+\tau} q_i \rangle, \tag{9.38}$$

которые можно интегрировать как средние значения совместных функций рапределения

$$\langle q_{i+\tau} q_i \rangle = \iint dq_{i+\tau} dq_i\, q_{i+\tau} q_i P(q_{i+\tau}, q_i). \tag{9.39}$$

Наша задача состоит в том, чтобы получить несмещенные оценки совместных функций распределения, входящих в формулу (9.39), вместо условных вероятностей, изучением которых мы занимались раньше. Для упрощения задачи предположим, что измерения производятся в стационарных условиях. Это позволяет нам представить совместную вероятность в виде произведения условной вероятности и стационарного распределения вероятности

$$P(q_{i+\tau}, q_i) = P(q_{i+\tau} | q_i) P_{\text{стац}}(q_i). \tag{9.40}$$

Поскольку относительно q_i мы располагаем не точной информацией, а лишь информацией, содержащейся в корреляционных функциях, нам понадобятся дополнительные ограничения. Мы хотим воспользоваться следующими ограничениями:

$$\langle q_i^\mu \rangle, \qquad \mu = 1, \ldots, m; \tag{9.41}$$

$$\langle q_{i+\tau} q_i^\nu \rangle, \qquad \nu = 0, \ldots, n; \tag{9.42}$$

$$\langle q_{i+\tau}^2 q_i^\varkappa \rangle, \qquad \varkappa = 0, \ldots, k, \tag{9.43}$$

где числа m, n и k удовлетворяют некоторым условиям самосогласованности, о которых пойдет речь дальше, и могут быть выбраны так, чтобы обеспечить разумную сходимость данных. Можно ожидать, что вблизи неравновесных фазовых переходов числа n и k будут ограничены значениями $n = k = 3$. Между разд. 9.3 и предшествующим разделом существует фундаментальное различие. Как нетрудно видеть, множители λ_i теперь не зависят от q_i. Поскольку принцип максимума информационной энтропии, или, точнее говоря, принцип максимума калибра, относится к двум переменным $q_{i+\tau}$ и q_i, принцип максимума информационной энтропии приводит к формуле

$$P_{\text{стац}} = \exp\left(\lambda_{\text{стац}} + \sum_{\mu=1}^{m} \lambda_{\mu,\,\text{стац}} q_i^\mu\right), \tag{9.44}$$

в то время как принцип максимума калибра приводит к следующему выражению для совместной вероятности:

$$P(q_{i+\tau}, q_i) = \exp\left[\lambda_0 + a(q_i) + q_{i+\tau} b(q_i) + q_{i+\tau}^2 C(q_i)\right]. \tag{9.45}$$

Здесь мы воспользовались для краткости следующими обозначениями:

$$a(q_i) = \sum_{\mu=1}^{m} \lambda_\mu^{(1)} q_i^\mu, \tag{9.46}$$

$$b(q_i) = \sum_{\nu=0}^{n} \lambda_\nu^{(2)} q_i^\nu, \tag{9.47}$$

$$C(q_i) = \sum_{\varkappa=0}^{k} \lambda_\varkappa^{(3)} q_i^{(\varkappa)}. \tag{9.48}$$

Соотношение (9.40) позволяет определить, что входящая в него условная вероятность определяется выражением

$$P(q_{i+\tau} | q_i) = \exp\left[\hat{\lambda} + \hat{a}(q_i) + q_{i+\tau} b(q_i) + q_{i+\tau}^2 C(q_i)\right], \tag{9.49}$$

где

$$\hat{\lambda} = \lambda_0 - \lambda_{\text{стац}}, \tag{9.50}$$

$$\hat{a} = a - \sum_{\mu=1}^{m} \lambda_{\mu,\,\text{стац}} q_i^{\mu}. \tag{9.51}$$

Так мы нашли выражение для условной вероятности, в котором можно усмотреть прямую аналогию с выражением (9.8). Далее мы можем поступить так же, как в разд. 9.1, и установить следующие соответствия:

$$\lambda(q_i) \Leftrightarrow \hat{\lambda} + \hat{a}(q_i), \tag{9.52}$$

$$\lambda_1(q_i) \Leftrightarrow b(q_i), \tag{9.53}$$

$$\lambda_2(q_i) \Leftrightarrow C(q_i). \tag{9.54}$$

Правые части соотношений (9.52)–(9.54) должны удовлетворять некоторым условиям самосогласованности, так как λ зависит от λ_1 и λ_2 (см. соотношение (9.13)). Эта зависимость становится более сложной, если параметр λ_2 зависит от q_i. Так, соотношение (9.13) требует, по крайней мере в принципе, в качестве ограничений бесконечно высоких степеней величины q_i, или, иначе говоря, бесконечно многих моментов.

Однако можно ожидать, что в практических случаях окажется вполне достаточно и нескольких первых моментов. С другой стороны, если λ_2 зависит от q_i, то λ_2–полином порядка $m = 2n$. Теперь уже совершенно ясно, каким образом результаты этого раздела, относящиеся к одной переменной, рассматриваемой в данный момент времени, могут быть обобщены на случай, когда в данный момент времени рассматривается вектор состояния. Из соображений экономии места мы не приводим здесь соответствующие результаты для векторного случая.

В этом разделе мы показали, каким образом простые моменты или корреляционные функции, если их использовать в качестве ограничений, позволяют вывести основные выражения статистической физики, а именно выражения для совместных вероятностей и интегралов по траекториям, которые в свою очередь могут быть интерпретированы как решения уравнения Фоккера–Планка рассматриваемой задачи. Таким образом, наш подход позволяет строить предположения относительно механизмов, кроющихся за наблюдаемыми экспериментальными данными; мы можем определять дрейфовые коэффициенты, характеризующие детерминистические силы, и коэффициенты диффузии, служащие мерой флуктуирующих сил. В приложениях не всегда возможно предполагать, что стохастический процесс, лежащий в основе рассматриваемой задачи, является марковским или непрерывным марковским процессом. В подобной ситуации наш подход позволяет определять действующие на рассматриваемую систему детерминистические и флуктуирующие силы. Ясно, что предлагаемый нами подход может быть обобщен в различных направлениях, например с учетом высших моментов или

корреляционных функций. Мы можем определять совместные функции распределения или условные вероятности на бóльших интервалах времени. С другой стороны, в подобных случаях наш метод может утратить присущее ему изящество.

9.4. Уравнение Фоккера–Планка, соответствующее пропагатору на малых временах

Вывод уравнения Фоккера–Планка, соответствующего условной вероятности (9.35),–хорошо известная процедура. Чтобы показать, к какому именно типу (Ито и Стратоновича) принадлежит уравнение Фоккера–Планка, соответствующее условной вероятности (9.32), кратко наметим основные этапы вычислений. Так как уравнение Фоккера–Планка получено в предположении бесконечно малых шагов по времени, достаточно рассмотреть два последовательных момента времени i и $i+\tau$.

Соответствующие функции распределения вероятности связаны между собой соотношением

$$P(\mathbf{q}_{i+\tau}) = \int P(\mathbf{q}_{i+\tau} \mid \mathbf{q}_i)\, P(\mathbf{q}_i)\, d^N q_i, \tag{9.55}$$

где условная вероятность определяется по формуле (9.32). Мы хотим вывести дифференциальное уравнение для $P(q_{i+\tau})$. Для этого, умножая соотношение (9.55) на произвольную функцию $g(\mathbf{q}_{i+\tau})$, получаем

$$A \equiv \int g(\mathbf{q}_{i+\tau})\, P(\mathbf{q}_{i+\tau})\, d^N q_{i+\tau} = \int d^N q_i\, P(\mathbf{q}_i) \underbrace{\int d^N q_{i+\tau}\, g(\mathbf{q}_{i+\tau})\, P(\mathbf{q}_{i+\tau} \mid \mathbf{q}_i)}_{F(q_i)}. \tag{9.56}$$

Чтобы вычислить функцию $F(\mathbf{q}_i)$, введем новый вектор переменных

$$\mathbf{q}_{i+\tau} - \mathbf{q}_i - \tau \mathbf{K}(\mathbf{q}_i) = \xi. \tag{9.57}$$

Тогда

$$F = \int d^N \xi \mathcal{N}(\mathbf{q}_i) \exp\left(-\xi \frac{G}{\tau} \xi\right) g\,[\mathbf{q}_i + \tau \mathbf{K}(\mathbf{q}_i) + \xi]. \tag{9.58}$$

Чтобы вычислить интеграл с точностью до линейных по τ членов, разложим функцию g в правой части соотношения (9.58) в степенной ряд по ξ до квадратичных членов и в степенной ряд по τ до членов первого порядка.

Используя соотношения

$$\int \mathcal{N}(\mathbf{q}_i) \exp\left(-\xi \frac{G}{\tau} \xi\right) d^N \xi = 1, \tag{9.59}$$

$$\int \mathcal{N}(\mathbf{q}_i) \exp\left(-\xi \frac{G}{\tau} \xi\right) \xi_k \xi_l\, d^N \xi = \frac{\tau}{2}(G^{-1})_{kl}, \tag{9.60}$$

а также то, что интегралы от нечетных степеней ξ равны нулю, получаем

$$F(\mathbf{q}_i) = g(\mathbf{q}_i) + \tau \mathbf{K}(\mathbf{q}_i) \nabla_{\mathbf{q}_i} g(\mathbf{q}_i) + \frac{\tau}{4} \sum_{k,l} (G^{-1})_{kl} \frac{\partial^2 g}{\partial q_{ik} \partial q_{il}}. \qquad (9.61)$$

Подставляя разложение (9.61) в соотношение (9.56) и интегрируя по частям, получаем (вместо (9.56)) соотношение

$$A = \int d^N q_i g(\mathbf{q}_i) \left\{ 1 - \tau \nabla_{\mathbf{q}_i} \mathbf{K}(\mathbf{q}_i) + \frac{\tau}{4} \sum_{k,l} \frac{\partial^2}{\partial q_{ik} \partial q_{il}} [G^{-1}(\mathbf{q}_i)]_{kl} \right\} P(\mathbf{q}_i). \qquad (9.62)$$

Введем обозначения

$$P(\mathbf{q}_{i+\tau}) \Rightarrow f(\mathbf{q};\ t + \tau), \qquad (9.63)$$

$$P(\mathbf{q}_i) \Rightarrow f(\mathbf{q};\ t). \qquad (9.64)$$

Так как функция g, входящая в правую часть соотношений (9.56) и (9.62), произвольна, соответствующие выражения должны быть равны и без интегрирования. Следовательно, должно выполняться соотношение

$$f(\mathbf{q};\ t + \tau) = f(\mathbf{q};\ t) + \tau L f(\mathbf{q};\ t), \qquad (9.65)$$

где L–оператор Фоккера–Планка

$$Lf = -\nabla_{\mathbf{q}} (K(\mathbf{q}) f) + \frac{1}{4} \sum_{k,l} \frac{\partial^2}{\partial q_k \partial q_l} ([G^{-1}(\mathbf{q})]_{kl} f), \qquad (9.66)$$

встречающийся в исчислении Ито. Таким образом, мы построили для уравнения Ито–Фоккера–Планка пропагатор для малых времен.

9.5. Можно ли вывести закон Ньютона из экспериментальных данных?

В этом разделе мы хотим обсудить возможность вывода с помощью предлагаемой нами процедуры (см. разд. 9.1, 9.2, 9.4) закона Ньютона

$$\dot{\mathbf{p}} = \mathbf{K}(\mathbf{q}) \qquad (9.67)$$

из экспериментальных данных относительно движения конкретных тел (например, планет) и силы **K**, порождающей это движение.

Прежде всего ясно, что мы должны рассматривать предельный случай, когда шум исчезает. Здесь условная вероятность представима в виде δ-функции, поэтому в одномерном случае

$$P(q_{i+\tau} | q_i) = \delta\left(\frac{q_{i+\tau} - q_i}{\tau} - K(q_i) \right), \qquad (9.68)$$

или, в случае многомерного обобщения,

$$P(\mathbf{q}_{i+\tau} \mid \mathbf{q}_i) = \delta\left(\frac{\mathbf{q}_{i+\tau} - q_i}{\tau} - \mathbf{K}(\mathbf{q}_i)\right). \tag{9.69}$$

На первый взгляд может показаться, что мы получили не закон Ньютона (9.67), а некий другой закон:

$$\dot{\mathbf{q}} = \mathbf{K}(\mathbf{q}). \tag{9.70}$$

Такие законы действительно были известны в науке еще до возникновения ньютоновской механики и описывают движение с сильным затуханием. Но мы легко можем восстановить закон Ньютона, если включим в множество наблюдаемых переменных $\mathbf{q}$ не только пространственные координаты частиц (или небесных тел), *но и их импульсы (или скорости)*. Нетрудно выяснить, как выглядят в этом случае (когда мы предполагаем ньютоновскую механику) условные моменты (9.5) и (9.6), и убедиться, что в действительности мы получаем выражение (9.67).

Можно надеяться, что распространение нашей процедуры на квантовую механику позволит вывести фундаментальные законы физики элементарных систем.

Упражнение. Исходя из закона Ньютона $\dot{p} = K(q)$, вычислите условные моменты (9.21) и (9.22):

$$\langle q_{l,\, i+\tau}\rangle_{q_{k,\, i}}, \qquad \langle q^2_{l,\, i+\tau}\rangle_{q_{k,\, i}}, \qquad l = 1,\, 2; \qquad k = 1,\, 2$$

при

$$q_1 \equiv q, \qquad q_2 \equiv p$$

и бесконечном малом τ. Воспользуйтесь изложенным выше формализмом для вывода соотношения (9.69).

Глава 10

Приложение к некоторым физическим системам

В настоящее время известно большое число неравновесных фазовых переходов, происходящих в самых различных системах. В разд. 5.2 и 5.3 мы продемонстрировали новый подход на примере одно- и многомодового лазера (без фазовых соотношений). В этом разделе мы сначала продолжим исследование лазеров, обращая внимание на более сложные случаи, а затем перейдем к рассмотрению конвективной неустойчивости в гидродинамике.

10.1. Мультимодовые лазеры с фазовыми соотношениями

Рассмотрим кольцевой лазер, в котором напряженность электрического поля E может быть записана в виде

$$E(x, t) = \sum_l B_l(t) \exp(\mathrm{i}k_l x - \mathrm{i}\omega_l t) + \text{к. с.}, \tag{10.1}$$

где l – индекс мод в полости, а $B_l(t)$ – их зависящие от времени амплитуды, имеющие физическую размерность напряженности электрического поля.

Мы предполагаем, что волны распространяются только в одном направлении и что значения k и ω выбраны из некоторого интервала, малого по сравнению соответственно с центральным волновым числом k_0 и центральной частотой ω_0:

$$k_l \in k_0 \pm \Delta k, \qquad \omega_l \in \omega_0 \pm \Delta\omega. \tag{10.2}$$

Рассмотрим пространственное и временное среднее напряженности электрического поля E и ее степеней (временное среднее определим так же, как в гл. 5). Нетрудно видеть, что

$$\overline{E} = 0, \tag{10.3}$$

$$\overline{E_l(x, t) E_{l'}(x, t)} = 2\delta_{ll'} B_l B_{l'}^*, \tag{10.4}$$

$$\overline{E_l E_{l'} E_{l''}} = 0. \tag{10.5}$$

Интегрируя по плоским волнам и используя правила отбора волновых чисел k, получаем

$$\overline{E_l E_{l'} E_{l''} E_{l'''}} = C \sum_{\pm} \delta(\pm k_l \pm k_{l'} \pm k_{l''} \pm k_{l'''}) B_l^{\pm} B_{l'}^{\pm} B_{l''}^{\pm} B_{l'''}^{\pm}, \tag{10.6}$$

где $B^+ = B^*$ и $B^- = B$.

Введем следующие обозначения (см. разд. 3.3):
индекс i величины p_i:

$$i = (B_1, B_1^*, B_2, B_2^*, \ldots, B_M, B_M^*), \tag{10.7}$$

ограничения:

$$f_l = \langle B_l^* B_l \rangle, \tag{10.8}$$

$$f_{l,l',l'',l'''} = \langle B_l^* B_{l'}^* B_{l''} B_{l'''} \rangle \delta(k_l + k_{l'} - k_{l''} - k_{l'''}). \tag{10.9}$$

Применение принципа максимума информационной энтропии приводит к окончательному результату:

$$P(\mathbf{B}, \mathbf{B}^*) = \exp\Big[-\lambda - \sum_l \lambda_l B_l^* B_l - \sum_{l,l',l'',l'''} \lambda_{l\,l'\,l''\,l'''} B_l^* B_{l'}^* B_{l''} B_{l'''} \times \times \delta(k_l + k_{l'} - k_{l''} - k_{l'''}\Big]. \tag{10.10}$$

Ранее этот результат был получен в рамках микроскопической теории для случая, когда выполняется принцип детального равновесия.

10.2. Одномодовый лазер с учетом поляризации и инверсии

В этом разделе мы рассмотрим одномодовый лазер, в описании которого поляризация и инверсия атомной системы включены в качестве динамических переменных. Для простоты нас будут интересовать бегущие волны. В этом случае мы имеем следующие переменные:

$$E(x, t) = B(t)\exp(\mathrm{i}kx - \mathrm{i}\omega t) + \text{к.с. поле}, \tag{10.11}$$

$$P(x, t) = P(t)\exp(\mathrm{i}kx - \mathrm{i}\omega t) + \text{к.с. поляризация}, \tag{10.12}$$

$$D(x, t) - \text{медленно изменяющаяся инверсия}, \tag{10.13}$$

где $B(t)$, $P(t)$ и D – медленно изменяющиеся функции времени. Производя усреднение по пространственным переменным, получаем

$$\overline{E} = \overline{P} = 0, \qquad \overline{D} = 0, \tag{10.14}$$

$$\overline{E^* E} = 2B^* B, \qquad \overline{E^* E^* E E} = 4B^* B^* B B, \tag{10.15}$$

$$\overline{|P(x, t)|^2} = 2P^*(t) P(t), \qquad \overline{P^* E} = 2P^*(t) B(t). \tag{10.16}$$

Обобщая предыдущий вариант нашего подхода, будем удерживать отличные от нуля члены до четвертого порядка включительно. Функция распределения принимает при этом следующий вид:

$$P = \exp(\lambda + \lambda_1 |E|^2 + \ldots + \lambda_{27} D^4). \tag{10.17}$$

К сожалению, не существует микроскопической теории, в которой функция распределения для рассматриваемой нами задачи была бы

вычислена точно. Поэтому мы, как и прежде, ограничимся адиабатическим приближением, в котором полностью сохраним функции распределения P и D. Как и прежде (см. разд. 7.11), мы исходим из уравнений для лазера

$$\dot{E} = -\varkappa E + gP + E, \tag{10.18}$$

$$\dot{P} = -\gamma P + gED + \Gamma, \tag{10.19}$$

$$\dot{D} = \gamma_{\parallel}(D_0 - D) - g(EP^* + E^*P) + \Gamma_D. \tag{10.20}$$

Вводя адиабатическое приближение

$$\dot{P} = \dot{D} \approx 0, \tag{10.21}$$

где мы включаем флуктуирующие силы F, Γ и Γ_D, получаем из уравнения (10.19) соотношение

$$\Gamma = \gamma P - gED. \tag{10.22}$$

Поскольку известно, что флуктуирующая сила Γ имеет гауссовское распределение, мы заключаем, что величина

$$\gamma P - gED \tag{10.23}$$

имеет такую же функцию распределения. Аналогичным образом из уравнения (10.20) мы находим функцию распределения для разности двух первых скобок в правой части этого уравнения. Наконец, исключая адиабатически P и D вблизи порога, вместо уравнения (10.18) мы получаем уравнение

$$\dot{E} = -\varkappa E + GE - CE|E|^2 + F_{\text{полн}}, \tag{10.24}$$

где G и C – постоянные.

Уравнение (10.24) легко может быть преобразовано в уравнение Фоккера–Планка, которое, как было показано выше, легко решается в стационарных случаях. Полная функция распределения для E, P и D является произведением функций распределения, соответствующих E, P и D, и имеет вид

$$f = e^{\lambda} \exp\left(a|E|^2 - b|E|^4\right) \cdot \exp\left\{-\left(\frac{|\gamma P - gED|^2}{Q}\right) - \left(\frac{[g(EP^* + E^*P) + \gamma_{\parallel}(D - D_0)]^2}{Q'}\right)\right\}. \tag{10.25}$$

При сравнении функции распределения (10.25) с предполагаемой функцией распределения (10.17) мы обнаруживаем, что некоторые множители Лагранжа λ обращаются в нуль (табл. 10.1). Вопрос о том, обращаются ли некоторые λ в нуль вследствие предложенной нами процедуры адиабатического приближения или они равны нулю всегда, даже в случае уравнений (10.18)–(10.20) общего вида, остается пока открытым.

Таблица 10.1. **Обращающиеся (/) и не обращающиеся (×) в нуль множители Лагранжа**

	1	$\|E\|^2$	$\|P\|^2$	EP^*	E^*P	D	D^2
1	×	×	×	×	×	×	×
$\|E\|^2$	×	×	×	/	/	/	×
$\|P\|^2$	×	×	/	/	/	/	/
EP^*	×	/	/	×	×	×	/
E^*P	×	/	/	×	×	×	/
D	×	/	/	×	×	×	/
D^2	×	×	/	/	/	/	/

10.3. Гидродинамика: конвективная неустойчивость

Прежде всего напомним читателю, что такое конвективная неустойчивость. Рассмотрим жидкость, которая находится в прямоугольном сосуде, подогреваемом снизу. Если разность температур между нижней и верхней поверхностью жидкости мала, то перенос тепла осуществляется за счет теплопроводности и никакого макроскопического движения при этом не наблюдается. Но, когда разность температур превышает некоторое критическое значение, в жидкости возникает макроскопическое движение: возникает система конвективных валов. Она характеризуется полем вертикальных скоростей с длиной волны λ, которая вдвое больше диаметра конвективного вала. Поскольку такую синусоидальную вариацию поля вертикальных скоростей можно рассматривать как пространственную структуру, напрашивается мысль выбрать в качестве параметра порядка амплитуду соответствующей синусоиды (или плоской волны). Используя ее моменты до четвертого порядка включительно и замечая, что по соображениям симметрии моменты нечетного порядка должны быть равны нулю, мы сразу же получаем функцию распределения

$$f(\xi) = N \exp(\lambda\xi^2 - \beta\xi^4), \qquad (10.26)$$

которая согласуется с функцией распределения, найденной в рамках микроскопической теории.

Если предположить, что нам удалось измерить первые четыре пространственные корреляционные функции поля скоростей, то можно воспользоваться методом гл. 6 и определить совместную функцию распределения $P(\xi_u, \xi_s)$ в виде

$$P = f(\xi_u)\prod_s P(\xi_s \mid \xi_u). \tag{10.27}$$

Но, поскольку такой подход связан с довольно громоздкими вычислениями и не позволяет по-новому взглянуть на существо дела, мы не будем сейчас углубляться в эту проблему.

Рассмотрим вместо нее другое явление, например шестиугольные ячейки, образующиеся в некоторых жидкостях, налитых в круглый сосуд. Такие шестиугольные ячейки можно интерпретировать как суперпозицию плоских волн $\xi_k \exp[i\mathbf{k}\cdot\mathbf{x}]$, волновые векторы которых образуют равносторонний треугольник. Таким образом, в качестве параметров порядка ξ мы используем комплексные амплитуды этих плоских волн и отличаем их по индексам 0, $\pi/3$, $2\pi/3$, соответствующим их ориентации на плоскости, параллельной поверхности жидкости. Как показывают рассуждения, аналогичные приведенным выше (см. разд. 10.1) правилам отбора для мультимодового лазера, некоторые моменты обращаются в нуль. В результате остается только часть моментов и соответствующие им множители Лагранжа. Отличными от нуля оказываются следующие моменты:

$$\langle |\xi_0|^2 \rangle, \qquad \langle |\xi_{\pi/3}|^2 \rangle, \qquad \langle |\xi_{2\pi/3}|^2 \rangle, \tag{10.28}$$

$\langle \xi_0^* \, \xi_{\pi/3} \, \xi_{2\pi/3} \rangle$, моменты получающиеся из последнего при циклических перестановках, и, наконец,

$$\langle |\xi_j|^2 |\xi_k|^2 \rangle, \qquad j, k = 0, \quad \pi/3, \; 2\pi/3. \tag{10.29}$$

После этого уже не составляет труда построить функцию распределения. Она также согласуется с функцией распределения, получаемой в рамках микроскопической теории.

Глава 11

Переходы между паттернами поведения в биологии. Пример: движения рук

В этой главе мы хотим показать, как методы, изложенные в предыдущих главах, могут быть применены к биологическим системам. Приводимые нами примеры позволят особенно ясно понять, сколь мощным оказывается предлагаемый нами подход при моделировании поведения сложных систем.

11.1. Некоторые экспериментальные факты

При исследовании непроизвольных осцилляторных движений двух указательных пальцев Келсо обнаружил интересное явление. По команде увеличить частоту противофазных антисимметричных движений указательных пальцев (активность разгибателей одного пальца сочетается с активностью сгибателями другого пальца) у подопытного субъекта внезапно происходил переход к синфазному симметричному движению обоих пальцев (одновременной активности гомологичных групп мышц обоих пальцев) (см. рис. 11.1, 11.2).

Обнаруженное Келсо явление наблюдается не только при движении пальцев. В более поздней работе, используя аналогичные экспериментальные манипуляции, исследователь обнаружил аналогичные переходы в движениях кистей рук относительно запястий: антисимметричные противофазные движения внезапно сменялись симметричными синфазными движениями. Хотя у различных субъектов этот фазовый переход происходил при сильно отличающихся частотах, само явление тем не менее было вполне предсказуемо. Если же частоту переходов измерять в единицах предпочтительной частоты (т. е. ввести для каждого субъекта независимую меру частоты движений, при которой он (или она) готов совершать циклически повто-

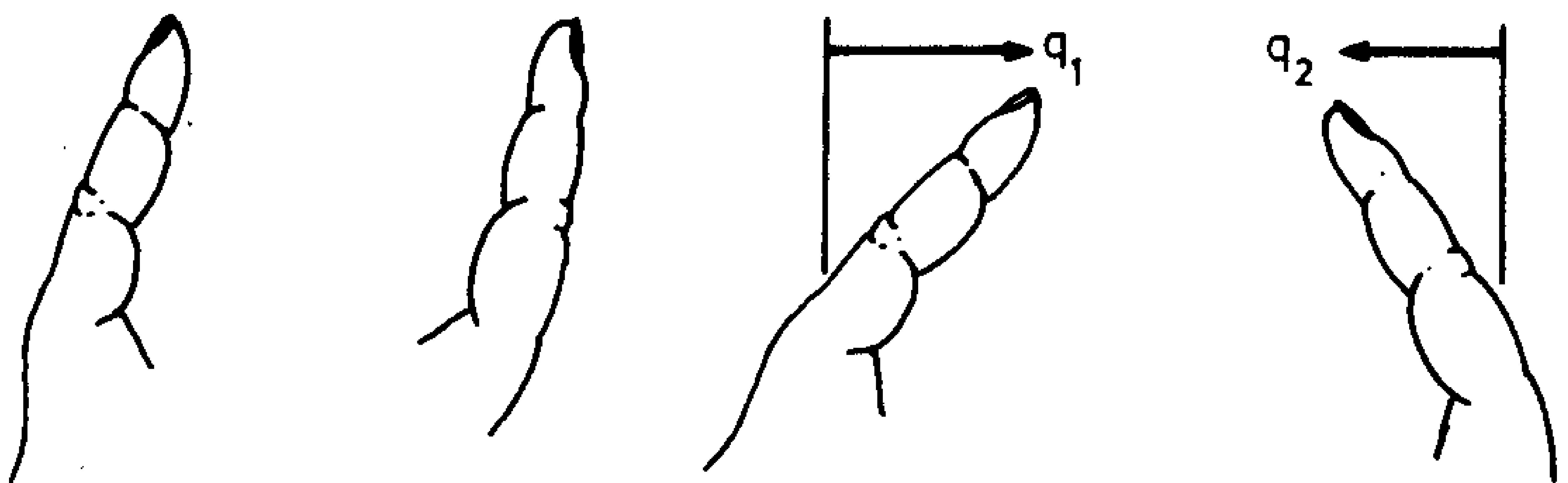
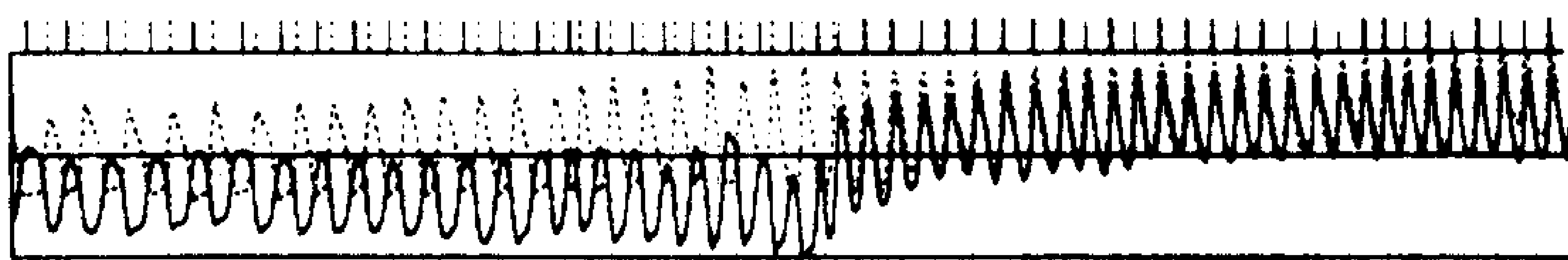


Рис. 11.1. Непроизвольный переход от параллельного движения пальцев (слева) к антипараллельному симметричному движению (справа).



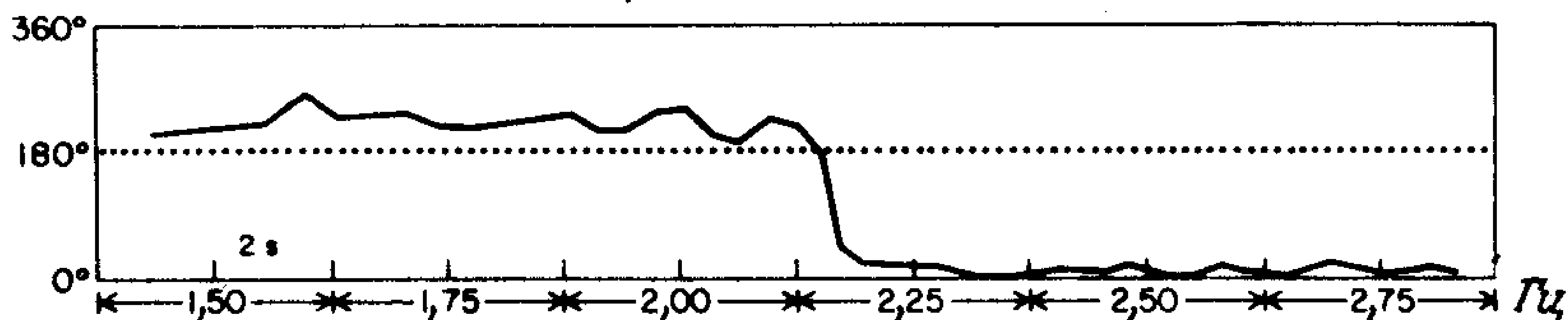


Рис. 11.2. Вверху: движения во времени левой (сплошная линия) и правой (пунктирная линия) рук. Субъект по команде экспериментатора просто увеличивает циклическую частоту в антисимметричной моде. Внизу: рассогласование по фазе между двумя руками. Пики движений одной руки образуют массив данных, подлежащих обработке. Для каждого пика вычисляется его сдвиг по фазе относительно периода от пика до пика другого массива. Построенный график повторяет фазовую кривую, на нем отчетливо видны и отставания, и опережения по фазе.

ряющиеся движения «хоть целый день»), то получающееся безразмерное отношение, или «критическое значение», оказывалось одинаковым у всех подопытных субъектов. Введение сопротивления трения движению сопровождалось систематическим изменением предпочтительной частоты и частоты перехода для каждого субъекта, но не изменяло критического значения для всех субъектов.

Наиболее поразительная особенность этих простых экспериментов состоит во внезапном и совершенно непроизвольном изменении упорядочения, или фазировки, активности мышечных групп при некоторой критической, определяемой индивидуальными особенностями субъекта частоте. В этом отношении эксперимент Келсо напоминает эксперименты по переходу с одного типа движения на другой в экспериментах по изучению двигательной активности. Например, Шик и др. показали, что при монотонном усилении электрической стимуляции среднего мозга у децеребрированной кошки наблюдалось не только увеличение частоты двигательной активности, но и (при превышении некоторого порога) смена типа движения. Так же, как в экспериментах по изучению движений указательных пальцев и кистей рук, в экспериментах по изучению двигательной активности наблюдался внезапный «перескок» от одной фазы к другой. В этих же экспериментах были обнаружены неустойчивые области, в которых кошка переходила с «рыси» на «галоп» и с «галопа» на «рысь». Хотя данные по изучению движений указатель-

ных пальцев и кистей рук, а также данные по движению конечностей у четвероногих убедительно свидетельствуют о том, что изменения в координации движений могут быть вызваны изменениями одного параметра, процессы, происходящие в нервных клетках, которыми обусловлены переходы в фазах моторной активности, пока еще мало понятны.

Резюмируя, можно утверждать, что основные особенности описанных выше экспериментов сводятся к следующим: (1) наличию только двух устойчивых фазовых состояний («аттракторов») в движениях рук (какая именно фаза наблюдается, зависит от того, как приготовлена система, т. е. от команды двигать кистями или указательными пальцами синфазно или в противофазе); (2) резкому переходу из одного состояния-аттрактора в другое при некоторой критической частоте циклических движений; (3) существованию в закритической области (при частоте выше критической) только одной моды (симметричных синфазных движений); (4) сохранению симметричной моды при понижении циклической частоты (иначе говоря, при уменьшении частоты система не возвращается в первоначально приготовленное состояние, что свидетельствует о сосуществовании областей притяжения симметричной и антисимметричной мод и подавлении одной из них).

11.2. Как моделировать переход

Попытаемся теперь с помощью изложенных выше методов построить модели экспериментально наблюдавшихся данных, о которых мы только что говорили. Обозначим через q_1 и q_2, как показано на рис. 11.1, перемещения концов указательных пальцев. Тогда

$$q_1 = r_1 \cos(\omega t + \phi_1), \tag{11.1}$$

$$q_2 = r_2 \cos(\omega t + \phi_2), \tag{11.2}$$

где r_1 и r_2 – амплитуды, которые, как показывают экспериментальные данные, можно считать не зависящими от времени, но зависящими от частоты ω, с которой должны двигаться пальцы. Фазы ϕ_1 и ϕ_2 – медленно изменяющиеся функции времени.

В дальнейшем мы будем предполагать, что

$$r_1 = r_2 = r. \tag{11.3}$$

Чтобы применить наш общий подход, мы рассматриваем q_1 и q_2 как наблюдаемые величины, моменты которых, например

$$\langle q_j \rangle \quad \text{или} \quad \langle q_j q_l \rangle, \tag{11.4}$$

могут быть определены. Следуя изложенному выше варианту нашего подхода, мы будем учитывать моменты до четвертого порядка включительно. По аналогии с лазером нас будут интересовать не относительно быстрые осцилляции, а медленно изменяющиеся фазы

ϕ_1 и ϕ_2. Начнем с среднего по времени, определяемого соотношения

$$\langle q_j \rangle_t = \frac{1}{T}\int_0^T q_j\,dt, \tag{11.5}$$

где

$$T = \frac{2\pi}{\omega}. \tag{11.6}$$

Поскольку, как показывают эксперименты, ϕ_1 и ϕ_2 – медленно изменяющиеся функции времени, мы можем воспользоваться соотношением

$$\frac{1}{T}\int_0^T \cos(\omega t + \phi_j)\,dt = 0, \tag{11.7}$$

или, если косинус записать с помощью экспоненты,

$$\frac{1}{T}\int_0^T \exp(\mathrm{i}\omega t + \mathrm{i}\phi_j)\,dt = 0. \tag{11.8}$$

С учетом соотношения (11.8) мы можем потребовать, чтобы выполнялось условие

$$\frac{1}{T}\int_0^T \exp[\mathrm{i}n(\omega t + \phi_j)]\,dt = 0, \tag{11.9}$$

где n – целое число. Из соотношения (11.7) получаем

$$\frac{1}{T}\int_0^T q_j\,dt = 0. \tag{11.10}$$

Рассмотрим теперь вторые моменты

$$\frac{1}{T}\int_0^T q_j q_k\,dt, \tag{11.11}$$

где мы воспользовались представлением

$$q_j = \frac{r}{2}\{\exp[\mathrm{i}(\omega t + \phi_j)] + \exp[-\mathrm{i}(\omega t + \phi_j)]\}. \tag{11.12}$$

Рассмотрим различные члены, которые получаются при подстановке выражения (11.12) в формулу (11.11). Член

$$\frac{1}{T}\int_0^T \exp(2\mathrm{i}\omega t + \mathrm{i}\phi_j + \mathrm{i}\phi_k)\,dt \tag{11.13}$$

равен нулю, как это следует из условия (11.9). С другой стороны, член

$$\frac{1}{T}\int_0^T \exp[\mathrm{i}(\phi_j - \phi_k)]\,dt = \exp[\mathrm{i}(\phi_j - \phi_k)] \tag{11.14}$$

не содержит быстро осциллирующих частей, что позволяет нам заменить левую часть равенства (11.14) его правой частью. Объединяя формулы (11.13) и (11.14), получаем

$$\frac{1}{T}\int_0^T q_j q_k \, dt = \frac{r^2}{2}\cos(\phi_j - \phi_k). \tag{11.15}$$

Мы будем различать два случая.

а) $j = k$. Из соотношения (11.15) следует, что левая часть его является константой, не зависящей от времени:

$$(11.15) = \text{const}. \tag{11.16}$$

Поскольку такой член не содержит переменных, мы можем исключить его из дальнейшего рассмотрения.

б) $j \neq k$. Соотношение (11.15) наводит на мысль о целесообразности введения нового обозначения

$$\phi_2 - \phi_1 = \phi, \tag{11.17}$$

позволяющего записать его более кратко:

$$(11.15) = \frac{r^2}{2}\cos\phi. \tag{11.18}$$

Рассмотрим теперь моменты третьего порядка

$$\frac{1}{T}\int_0^T q_{j_1} q_{j_2} q_{j_3} \, dt = 0. \tag{11.19}$$

Равенство (11.19) получается сразу же, как только мы примем во внимание, что в его левую часть могут входить только выражения вида

$$\exp(\mathrm{i}\omega t), \qquad \exp(3\mathrm{i}\omega t) \tag{11.20}$$

или комплексно-сопряженные с ними, которые, согласно (11.9), дают нулевой вклад. Рассмотрим, наконец, моменты четвертого порядка

$$\frac{1}{T}\int_0^T q_{j_1} q_{j_2} q_{j_3} q_{j_4} \, dt. \tag{11.21}$$

Подставляя представление (11.12) в формулу (11.21), получаем выражения, содержащие величины

$$\exp(4\mathrm{i}\omega t), \qquad \exp(2\mathrm{i}\omega t) \tag{11.22}$$

или комплексно-сопряженные с ними. При усреднении по T эти выражения обратятся в нуль. Но поскольку мы, наконец, получаем выражения, не содержащие быстро осциллирующих членов, то нам необходимо рассматривать только выражения вида

$$\exp[\mathrm{i}(\phi_{j_1} + \phi_{j_2} - \phi_{j_3} - \phi_{j_4})] \tag{11.23}$$

или выражение, получающееся из экспоненты (11.23) при любой перестановке индексов $j_1 \dots j_4$.

Выясним, какого рода выражения возникают, если индексам $j_1, \dots, j_4$ придавать значения 1 или 2. Перебирая все возможные случаи, получаем результаты, представленные в табл. 11.1.

Используя эти результаты, мы можем записать, что

$$\langle q_{j_1} q_{j_2} q_{j_3} q_{j_4} \rangle = C_1 + C_2 \cos\phi + C_3 \cos(2\phi). \tag{11.24}$$

Коэффициенты C_1, C_2 и C_3 пропорциональны r^4. Так как C_1 не содержит ни одной переменной, описывающей систему он не представляет для нас никакого интереса. Второй член учтен нами при рассмотрении вторых моментов, и поэтому его можно вычесть из соотношения (11.24). Таким образом, единственное существенное ограничение содержится в последнем члене соотношения (11.24). Резюмируя полученные результаты, можно утверждать, что из моментов второго и четвертого порядка следуют ограничения

$$r^2 \langle \cos\phi \rangle = f_1, \tag{11.25}$$

$$r^2 \langle \cos(2\phi) \rangle = f_2. \tag{11.26}$$

Таблица 11.1. Комбинации индексов j и соответствующие экспоненты

j_1	j_2	j_3	j_4	Аргумент экспоненты	
1	1	1	1	0	
2	2	2	2	0	
1	2	1	2	0	
1	2	2	1	0	
2	1	2	1	0	
2	1	1	2	0	
1	1	2	1	$\phi_1-\phi_2$	$= -\phi$
1	1	1	2	$\phi_1-\phi_2$	
2	1	2	2	$\phi_1-\phi_2$	
1	2	2	2	$\phi_1-\phi_2$	
2	1	1	1	$\phi_2-\phi_1$	$= \phi$
1	2	1	1	$\phi_2-\phi_1$	
2	2	1	2	$\phi_2-\phi_1$	
2	2	2	1	$\phi_2-\phi_1$	
1	1	2	2	$2(\phi_1 - \phi_2) = -2\phi$	
2	2	1	1	$2(\phi_2 - \phi_1) = 2\phi$	

Теперь уже несложно применить принцип максимума информационной энтропии и получить функцию распределения P следующего вида:

$$P(\phi) = \exp[-\lambda - \lambda_1 \cos\phi - \lambda_2 \cos(2\phi)]. \qquad (11.27)$$

Множители Лагранжа λ, λ_1 и λ_2 зависят от ω не только потому, что от ω зависят множители r^2 и r^4, стоящие перед средними значениями в формулах (11.25) и (11.26). Как показывают эксперименты, средние значения также зависят от ω. Экспоненту, входящую в соотношение (11.27), можно записать с помощью потенциала:

$$\hat{V}(\phi) = \lambda_1 \cos\phi + \lambda_2 \cos(2\phi). \qquad (11.28)$$

Этот потенциал определяет устойчивые и неустойчивые состояния системы. Графики зависимости $\hat{V}$ от ϕ при различных отношениях λ_1 и λ_2 представлены на рис. 11.3. Совершенно очевидно, что на графике, помещенном в верхнем левом углу, существует локальный минимум, соответствующий параллельному движению указательных пальцев. Но, когда параметр λ_2/λ_1 убывает, этот минимум становится более плоским и в конце концов исчезает совсем, так что шарик, изображающий состояние системы, скатывается в более низкий минимум, соответствующий антипараллельному (или, иначе говоря, симметричному) движению пальцев, о котором мы упоминали выше. Единственное предположение состоит в том, что отношение λ_2/λ_1 зависит от вынуждающей частоты ω. Но на этом возможности нашей модели далеко не исчерпываются. Рассмотрим случай, когда вынуж-

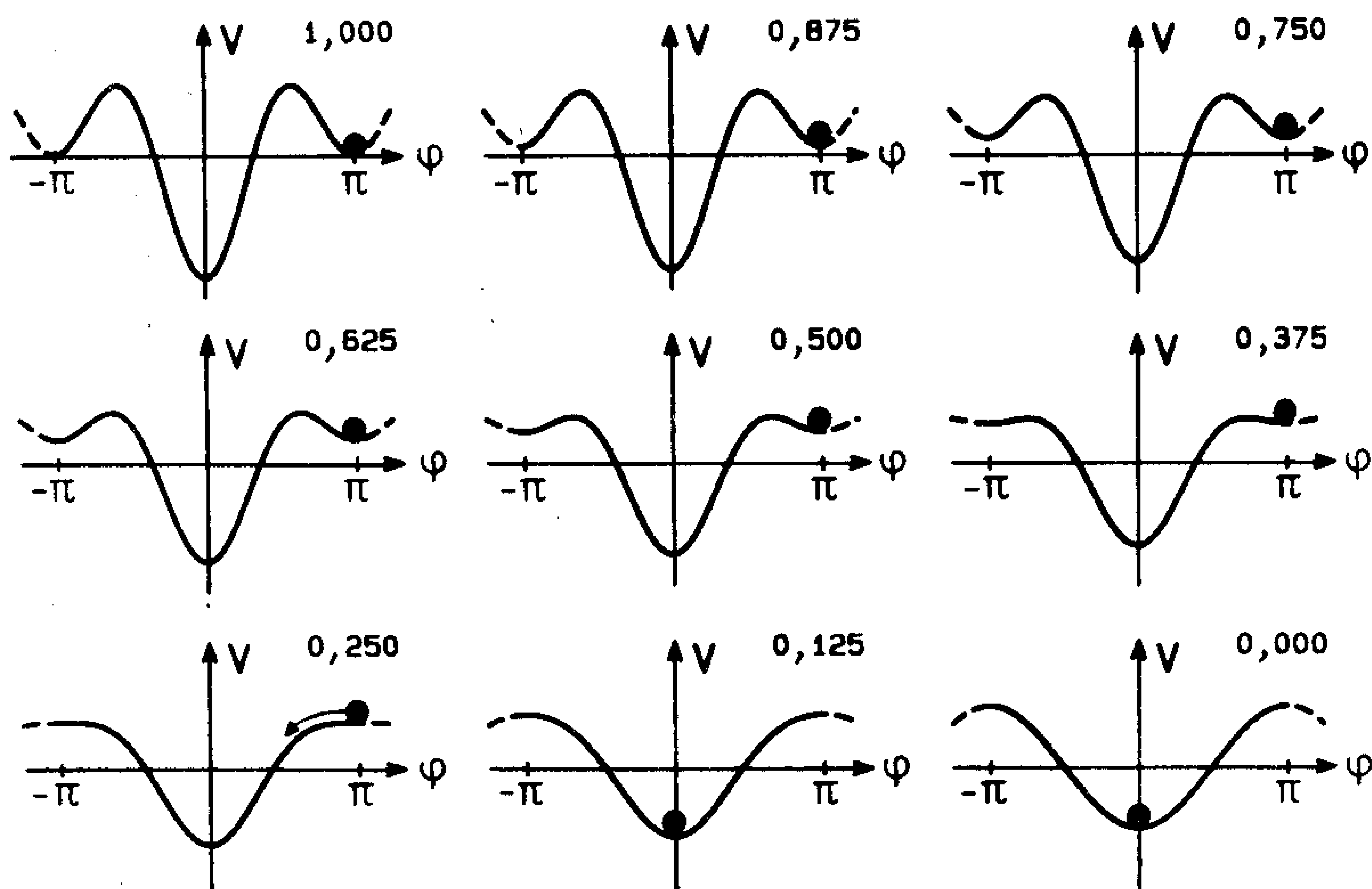


Рис. 11.3. Потенциал $\hat{V}$ при различных значениях λ_2/λ_1, указанных справа вверху у каждой кривой.

дающая частота настолько велика, что реализуется только состояние с симметричным движением пальцев. В этом случае (см. рис. 11.3) шарик, т. е. система, остается в состоянии с $\phi = 0$. Это означает, что подопытный субъект будет продолжать симметрично двигать указательными пальцами и после того, как частота уменьшится и станет ниже критической. Это теоретическое предсказание было проверено Келсо и подтвердилось.

Чтобы описать зависящие от времени процессы, в которых изменяется фаза ϕ, мы попытаемся записать уравнение Фоккера–Планка, которое соответствует фукции распределения (11.27). Как было показано в разд. 6.6, сделать это можно не единственным образом, поскольку источник шума может зависеть от переменных. Мы же примем простейшее из предположений и будем считать, что шум не зависит от ϕ. Тогда функцию распределения (11.27) можно сразу же отождествить с функцией распределения

$$P(\phi) = \mathcal{N} e^{-\hat{V}}, \qquad (11.29)$$

где мы для краткости ввели обозначение

$$\hat{V} = \frac{2}{Q} V \qquad (11.30)$$

и где Q – интенсивность источника шума. Из микроскопической теории (см. гл. 2) известно, что уравнение Фоккера–Планка имеет в этом случае следующий вид:

$$\dot{P}(\phi, t) = -\frac{\partial}{\partial \phi}[K(\phi) P] + \frac{Q \partial^2}{2 \partial \phi^2} P, \qquad (11.31)$$

где

$$K(\phi) = -\frac{\partial V}{\partial \phi}. \qquad (11.32)$$

В гл. 2 было показано, что уравнение Фоккера–Планка (11.31) соответствует уравнению Ланжевена

$$\dot{\phi} = K(\phi) + F(t), \qquad (11.33)$$

где правую часть с помощью соотношений (11.27), (11.28), (11.30), (11.32) и (11.33) можно привести к виду

$$\dot{\phi} = a \sin\phi + b \sin(2\phi) + F(t), \qquad (11.34)$$

здесь $a = Q\lambda_1/2$ и $b = Q\lambda_2$. Это уравнение позволяет сделать новые предсказания относительно характера движения указательных пальцев: мы можем решить уравнение (11.34) и тем самым вычислить корреляционные функции для фазы ϕ в различные моменты времени. Иначе говоря, мы получаем возможность исследовать с теорети-

ческой точки зрения переходные явления и сравнивать их с экспериментальными данными.

11.3. Критические флуктуации

Начнем с уравнения (11.34), которое мы выберем в качестве модельного уравнения. Оно содержит три параметра: a, b и неявно амплитуду флуктуирующей силы $F(t)$, которую мы запишем в виде $\sqrt{Q}\,\xi$, где ξ обладает следующими свойствами:

$$\langle \xi_t \rangle = 0, \tag{11.35}$$

$$\langle \xi_t \xi_{t'} \rangle = \delta(t - t'). \tag{11.36}$$

Прежде всего рассмотрим флуктуации фазы в симметричной моде. Так как в этом случае система близка к $\phi = 0$, мы можем линеаризовать уравнение Ланжевена (11.34), или, что эквивалентно, уравнение Фоккера–Планка (11.31). Например, уравнение Ланжевена принимает следующий вид:

$$\dot{\phi} = -(4b + a)\phi + \sqrt{Q}\,\xi_t. \tag{11.37}$$

Линеаризованное уравнение Фоккера–Планка решается без труда, и мы получаем возможность вычислить среднее значение абсолютной фазы:

$$\langle |\phi| \rangle_{\text{стац}} = \int\limits_{-\infty}^{+\infty} d\phi\, |\phi|\, P_{\text{стац}}(\phi) = \frac{1 - \exp(-\pi^2 d^2)}{\sqrt{\pi}\, d \operatorname{erf}\{\pi d\}}, \tag{11.38}$$

где $P_{\text{стац}}(\phi)$ – не зависящее от времени распределение вероятности для ϕ, и среднеквадратичное отклонение

$$\sigma_{\text{стац}} = \langle \phi^2 \rangle_{\text{стац}} - \langle |\phi| \rangle^2_{\text{стац}} = \frac{1}{2d^2} - \frac{\sqrt{\pi}\exp(-\pi^2 d^2)}{d \operatorname{erf}\{\pi d\}} - \langle |\Phi| \rangle^2_{\text{стац}}. \tag{11.39}$$

Совершенно ясно, что среднее значение (11.38) и среднеквадратичное отклонение (11.39) зависят от одного параметра d, который определяется выражением

$$d = \sqrt{\frac{4b + a}{Q}}. \tag{11.40}$$

Полученные результаты позволяют нам сравнивать экспериментальные данные относительно флуктуаций фазы с нашей моделью. Прежде всего обратимся к рис. 11.4, на котором представлены экспериментальные результаты.

На рис. 11.5 сплошными линиями показана зависимость среднего значения (11.38) и среднеквадратичного отклонения (11.39) от пара-

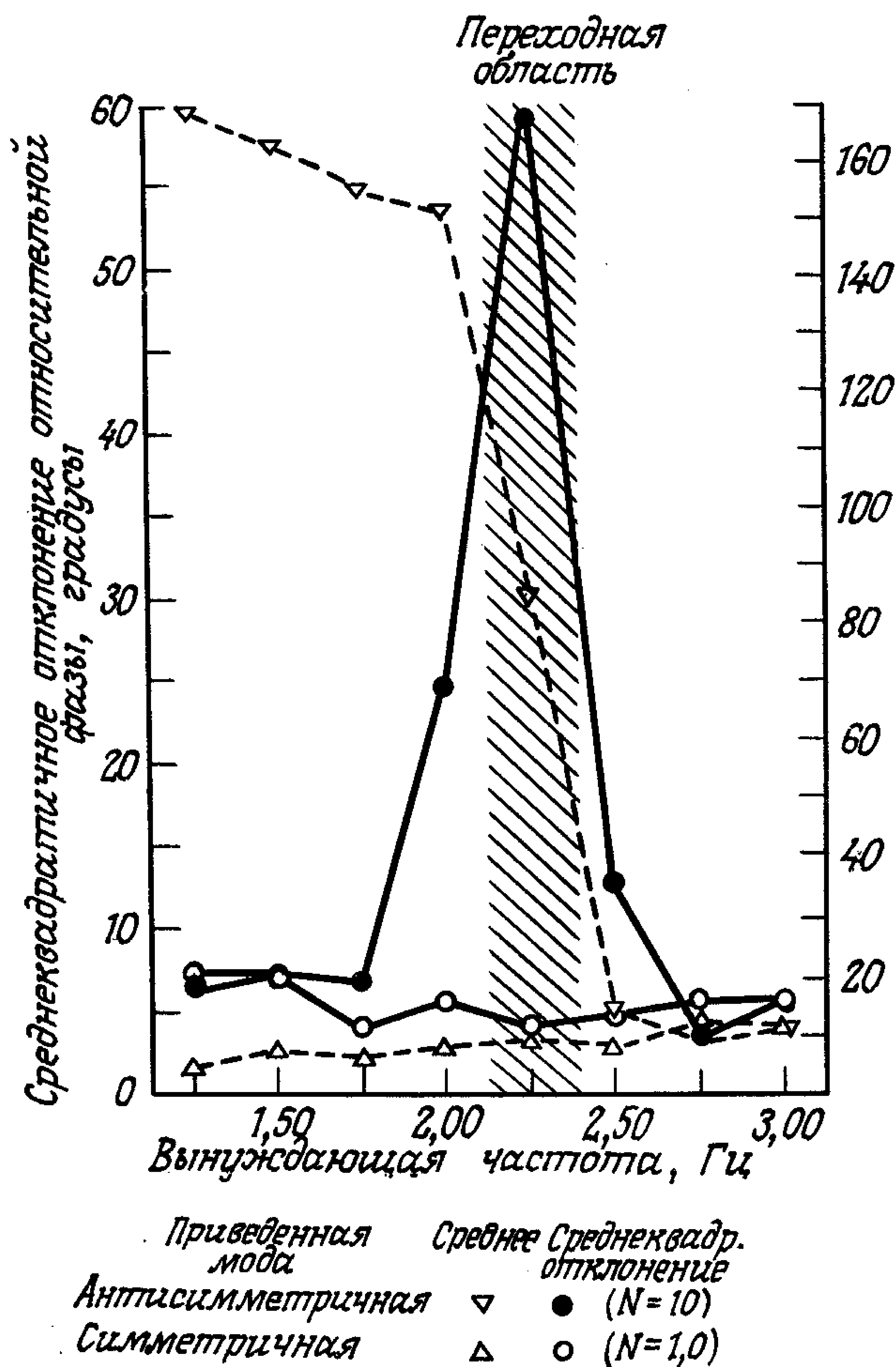


Рис. 11.4. Средний модуль фазы $\langle|\phi|\rangle$ и его среднеквадратичное отклонение $\{\langle\phi^2\rangle - \langle|\phi|^2\rangle\}$ определены в стационарном режиме – на плато по каждой частоте (т. е. в течение 3 с на каждой частоте). Каждая точка на графиках соответствует среднему по 10 повторениям эксперимента (из работы Келсо и Шольца, 1985).

метра d, а штриховыми линиями изображена зависимость тех величин, построенная по экспериментальным данным (например, по данным, приведенным на рис. 11.4). Это позволяет нам определить d. Действительно, как показывают экспериментальные данные,

$$\langle|\phi|\rangle \approx 8^\circ, \tag{11.41}$$

поэтому

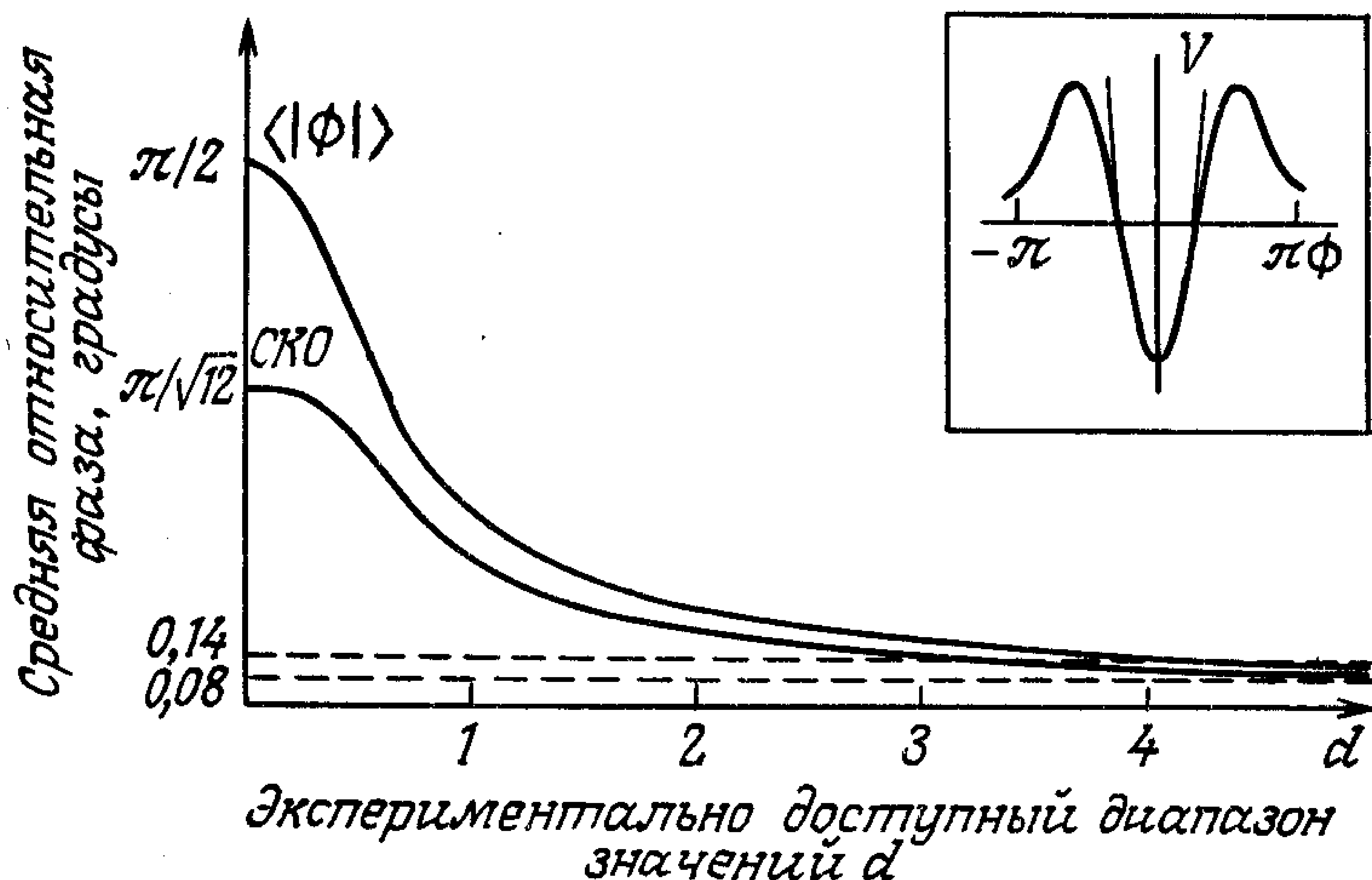


Рис. 11.5. Средний модуль фазы и среднеквадратичное отклонение, вычисленные в рамках локальной модели симметричной моды ($\phi_{\text{нач}} = 0$) как функции от $d = \{(a + 4b)/Q\}^{1/2}$. Две штриховые линии указывают экспериментальные значения для среднего ($= 0{,}14$) и среднеквадратичного ($= 0{,}08$), что позволяет указать экспериментально достижимый диапазон значений d. На врезке показан график потенциала (жирная линия) и локальной модели (тонкая линия) при $a = b = 1$ Гц.

$$3 \lesssim d \lesssim 5. \tag{11.42}$$

Из оценки

$$\sigma_{\text{стац}} \approx 4{,}5^\circ \tag{11.43}$$

мы заключаем, что

$$d \approx 3. \tag{11.44}$$

Совершенно ясно, что обеим экспериментальным оценкам можно удовлетворить, приняв значение

$$d \approx 4. \tag{11.45}$$

Время релаксации системы вблизи $\phi = 0$ нетрудно определить, решив линеаризованное уравнение Ланжевена в детерминистическом случае, т. е. без флуктуирующих сил. Мы получаем

$$\phi(t) = \phi(0) \exp[-(4b + a)t]. \tag{11.46}$$

Величина, обратная коэффициенту при t в показателе экспоненты, определяет время релаксации, поэтому по формуле (11.40) находим

$$\tau_{\text{рел}} = \frac{1}{4b + a} = \frac{1}{d^2 Q}. \tag{11.47}$$

Это время релаксации встречается во многих самых различных релаксационных явлениях и явлениях, связанных со скоростью реакции, например когда нам бывает необходимо перестать давить ногой на педаль газа при виде красного сигнала светофора. Грубая оценка дает $\tau_{\text{рел}} \approx 0{,}25$ с. Исходя из этого значения и $d \approx 4$, получаем из соотношения (11.47)

$$Q \approx 0{,}25 \text{ Гц}. \tag{11.48}$$

Тот же результат можно получить в рамках теории, учитывающей флуктуации и занимающейся исследованием корреляционной функции фазы ϕ в два различных момента времени. Проделаем такой же анализ для антисимметричной моды в случае, когда фаза ϕ центрирована на значении $+\pi$ или $-\pi$. Для этого введем малую величину ε, определяемую соотношением

$$\varepsilon = \begin{cases} \phi - \pi & \text{при} \quad 0 < \phi \leqslant \pi, \\ \phi + \pi & \text{при} \quad -\pi < \phi \leqslant 0. \end{cases} \tag{11.49}$$

Используя величину ε и удерживая в уравнении Ланжевена (11.37) только линейные члены, получаем

$$\dot{\varepsilon} = -(4b - a)\varepsilon + \sqrt{Q}\,\xi_t. \tag{11.50}$$

Нетрудно получить и соответствующее уравнение Фоккера–Планка. Его стационарное решение имеет вид

$$P_{\text{стац}}(\varepsilon) = \frac{f}{\sqrt{\pi}\,\text{erf}\{\pi f\}} e^{-f^2\varepsilon^2}, \tag{11.51}$$

где мы ввели параметр

$$f = \sqrt{\frac{4b - a}{Q}}. \tag{11.52}$$

Используя прежнее определение среднего значения фазы и стандартного отклонения, а также соотношение (11.51), получаем

$$\langle|\phi|\rangle_{\text{стац}} = \pi - \langle|\varepsilon|\rangle_{\text{стац}} = \pi - \frac{1 - \exp(-\pi^2 f^2)}{\sqrt{\pi} f\,\text{erf}\{\pi f\}}, \tag{11.53}$$

$$\sigma_{\text{стац}} = \langle\phi^2\rangle_{\text{стац}} - \langle|\phi|\rangle^2_{\text{стац}} = \langle\varepsilon^2\rangle_{\text{стац}} - \langle|\varepsilon|\rangle^2_{\text{стац}} =$$
$$= \frac{1}{2f^2} - \frac{\sqrt{\pi}\exp(-\pi^2 f^2)}{f\,\text{erf}\{\pi f\}} - \langle|\varepsilon|\rangle^2_{\text{стац}}. \tag{11.54}$$

Зависимость среднего значения (11.53) и среднеквадратичного отклонения (11.54) от параметра f представлена на рис. 11.6. Как видно из рис. 11.4, при каждом значении параметра f мы можем считывать вполне определенное среднее значение фазы ϕ и стандартного откло-

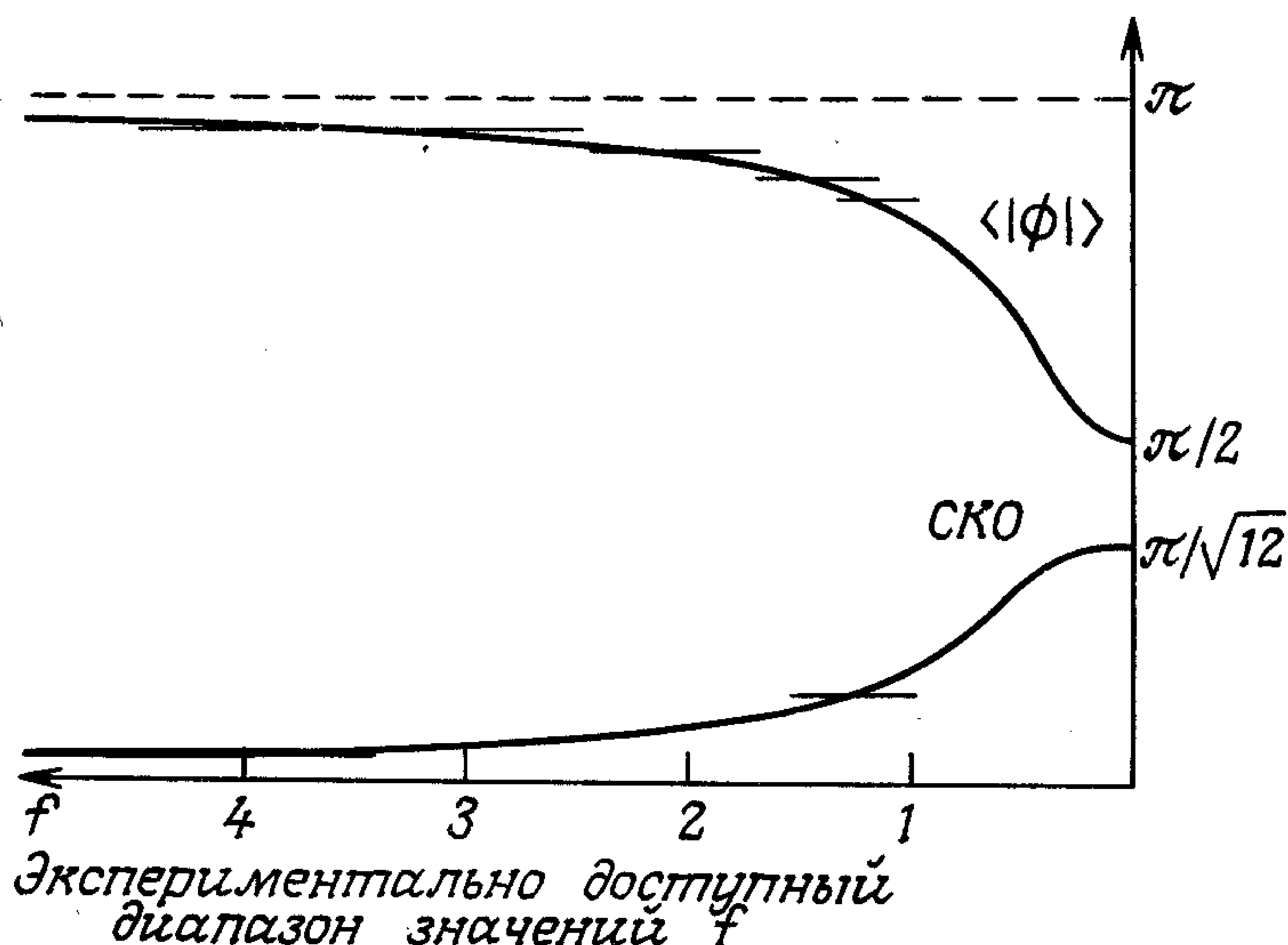


Рис. 11.6. Средний модуль фазы и его среднеквадратичное отклонение для локальной модели антисимметричной моды как функция параметра $f = [(4b - a)/Q]^{1/2}$. Приближение к переходу соответствует $f \to 0$. Для большей наглядности ось f направлена влево. Кривые были использованы для определения соответствующих значений f по экспериментальным значениям среднего $\langle|\phi|\rangle$ и его среднеквадратичного отклонения (СКО). Указан диапазон экспериментально допустимых значений f.

нения. Располагая такой парой значений, мы можем по графику на рис. 11.6 найти соответствующее значение f. Время релаксации по аналогии с формулой (11.47) можно определить как величину

$$\tau_{\text{рел}} = \frac{1}{4b - a} = \frac{1}{f^2 Q}. \tag{11.55}$$

Принимая снова значение $\tau_{\text{рел}} \approx 0{,}25$ с и выбирая значение параметра f в соответствии с экспериментальными данными в интервале от 1 до 4, мы получаем значение Q в интервале от 0,25 до 4, что по порядку величины согласуется с нашим предыдущим результатом для симметричной моды. Из соотношений (11.40) и (11.52) получаем

$$4b + a = Qd^2, \qquad 4b - a = Qf^2(\omega). \tag{11.56}$$

Так как относительно параметра f известно, что он зависит от частоты колебаний ω пальцев, мы можем получить из соотношений (11.56) частотную зависимость параметров a и b:

$$a(f) = \frac{Qd^2}{2} - \frac{Q}{2}f^2(\omega), \tag{11.57}$$

$$b(f) = \frac{Qd^2}{8} + \frac{Q}{8} f^2(\omega). \qquad (11.58)$$

Графики этих зависимостей представлены на рис. 11.7. Выбирая значение частоты, при котором происходит переход и соответствующее значение f, получаем следующие результаты: при $d \approx 4$, $Q = 0{,}25$ Гц оказывается, что $a_c = 4b_c$, поэтому приближенные значения этих параметров соответственно равны

$$a_c = 2{,}0 \text{ Гц}, \qquad (11.59)$$

$$b_c = 0{,}5 \text{ Гц}. \qquad (11.60)$$

Резюмируя, можно утверждать, что нам удалось получить разумные значения параметров и в каждом случае достичь согласия с четырьмя экспериментально определенными величинами, а именно с средними абсолютными фазами и стандартными отклонениями симметричной и антисимметричной мод с помощью трех модельных параметров. Разумеется, наша модель может быть усовершенствована, но я убежден, что и в существующем варианте она убедительно демонстрирует возможности использования предлагаемого нами подхода для моделирования сложных систем.

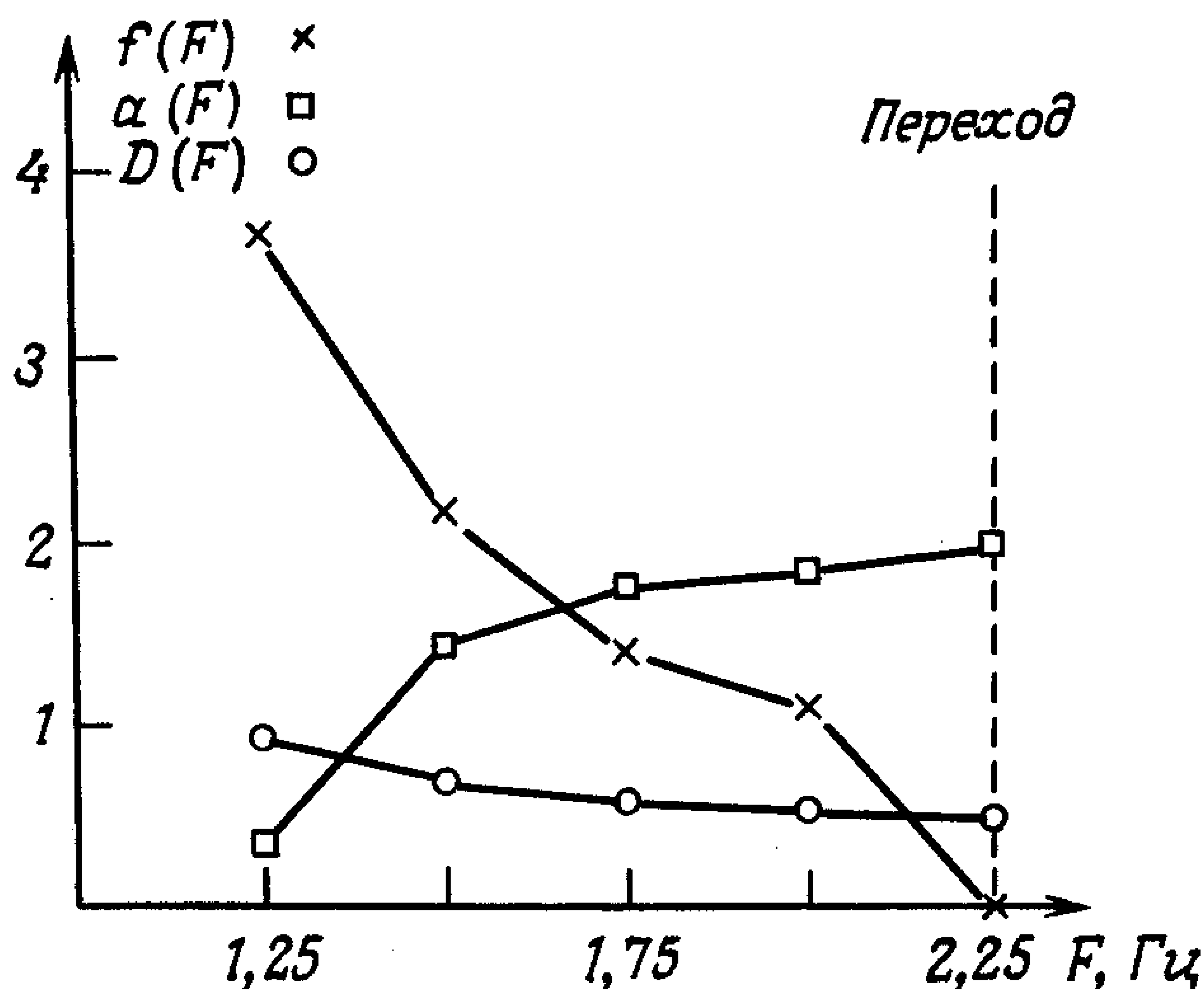


Рис. 11.7. Модельные параметры f, a и b как функции экспериментального управляющего параметра – частоты $F \equiv \omega$, определенные по данным о среднем и среднеквадратичном отклонении и по оценке времени релаксации.

11.4. Некоторые выводы

1) Моторная программа или самоорганизация?

В биологии для объяснения высокой координации мышц весьма часто прибегают к идее «моторной программы». Согласно этой идее, нейроны действуют более или менее как компьютер, в памяти которого хранится специальная программа. После обращения к этой программе нейрон начинает управлять отдельными движениями. При таком объяснении трудно понять, почему вообще должны возникать флуктуации: моторная программа вполне детерминирована и не допускает никаких флуктуаций. Вместе с тем критические флуктуации весьма характерны для неравновесных фазовых переходов, происходящих в тех случаях, когда имеет место самоорганизация. Именно поэтому мы считаем, что обнаруживаемые в таких экспериментах флуктуации существенно подкрепляют представление о мышцах и нейронах как об элементах, образующих самоорганизующуюся систему (аналогичную лазеру, в котором при изменении одного управляющего параметра могут происходить переходы между режимами различного типа). Вместе с тем эксперименты, о которых идет речь, не позволяют ответить на вопрос о том, происходит ли самоорганизация во всей системе «мышцы *и* нейроны» или только в нейронной подсистеме.

2) Сокращение информации

Высокая координация мышц и нейронов проявляется в специфических режимах макроскопического движения, описываемого в рассмотренном выше случае одним управляющим параметром. Именно с наличием одного-единственного управляющего параметра связано очень малое количество информации – малое по сравнению с информацией, необходимой для описания в отдельности состояний всех нейронов и мышц.

3) Морфогенез поведения

Как известно из многочисленных примеров из области физики и химии, самоорганизующиеся системы могут порождать некоторые характерные временны́е структуры. Мы видели, что определенные структуры, или паттерны, могут наблюдаться и в поведении биологических систем, например могут возникать характерные режимы в движении указательных пальцев. Было показано, что принцип максимума информационной энтропии позволяет находить адекватный параметр порядка и даже соответствующее ему уравнение. Можно надеяться, что и более сложные паттерны поведения могут быть описаны с помощью небольшого числа параметров порядка.

Глава 12
Распознавание образов

В этой главе мы хотим показать, как наш общий подход позволяет решить основную проблему распознавания образов. В разд. 12.1 мы выведем заново соотношения, позволяющие производить отбор характерных признаков, которые мы обсудим ниже. В разд. 12.2 мы обсудим принципиальную схему построения параллельного компьютера для распознавания образов. В последующих разделах (разд. 12.3 и 12.4) показано, что такая система может обучаться распознавать именно те образы, которые требуется выделить. Наконец, в разд. 12.5 наши результаты обобщаются на случай обучения распознаванию процессов. Разд. 12.1 носит несколько формальный характер, и читатели, интересующиеся лишь наиболее важными вопросами, затрагиваемыми в этой главе, могут ограничиться только вводной частью разд. 12.1 (до уравнения (12.1)), а затем переходить к следующим разделам.

12.1. Отбор характерных признаков

Поскольку мы не предполагаем у читателя предварительного знакомства с проблемой распознавания образов, начнем с краткого обзора. В этом разделе мы ограничимся распознаванием образов в последовательностях символов. Такими символами, в частности, могут быть буквы какого-нибудь алфавита. Мы рассмотрим особенно простой случай, когда алфавит состоит лишь из двух букв X и O (рис. 12.1, 12.2). Скобке, открытой слева, мы можем поставить в соответствие число 0, а скобке, открытой справа,– число 1. При таком кодировании букве X, как бы составленной из двух скобок) и (, соответствует комбинация (0, 1), а букве O, как бы составленной из двух скобок (и),– комбинация (1. 0).

В более общем плане мы можем разложить образ на характерные признаки и детали и поставить в соответствие каждому признаку символ q_j. При таком подходе всему образу соответствует вектор образа $\mathbf{q} = (q_1, q_2, \ldots, q_n)$ (рис. 12.3). Разумеется, вместо вектора мы можем рассматривать его конец – точку в многомерном пространстве. Следовательно, каждому образу мы ставим в соответствие вектор или точку. Компоненты этого вектора $\mathbf{q}$ не обязательно должны принимать значения 0 или 1, они могут быть и непрерывными величинами. Например, при разбиении образа на отдельные ячейки мы можем поставить в соответствие каждой точке j некоторую интенсивность q_j. На этой стадии можно утверждать, что распознава-

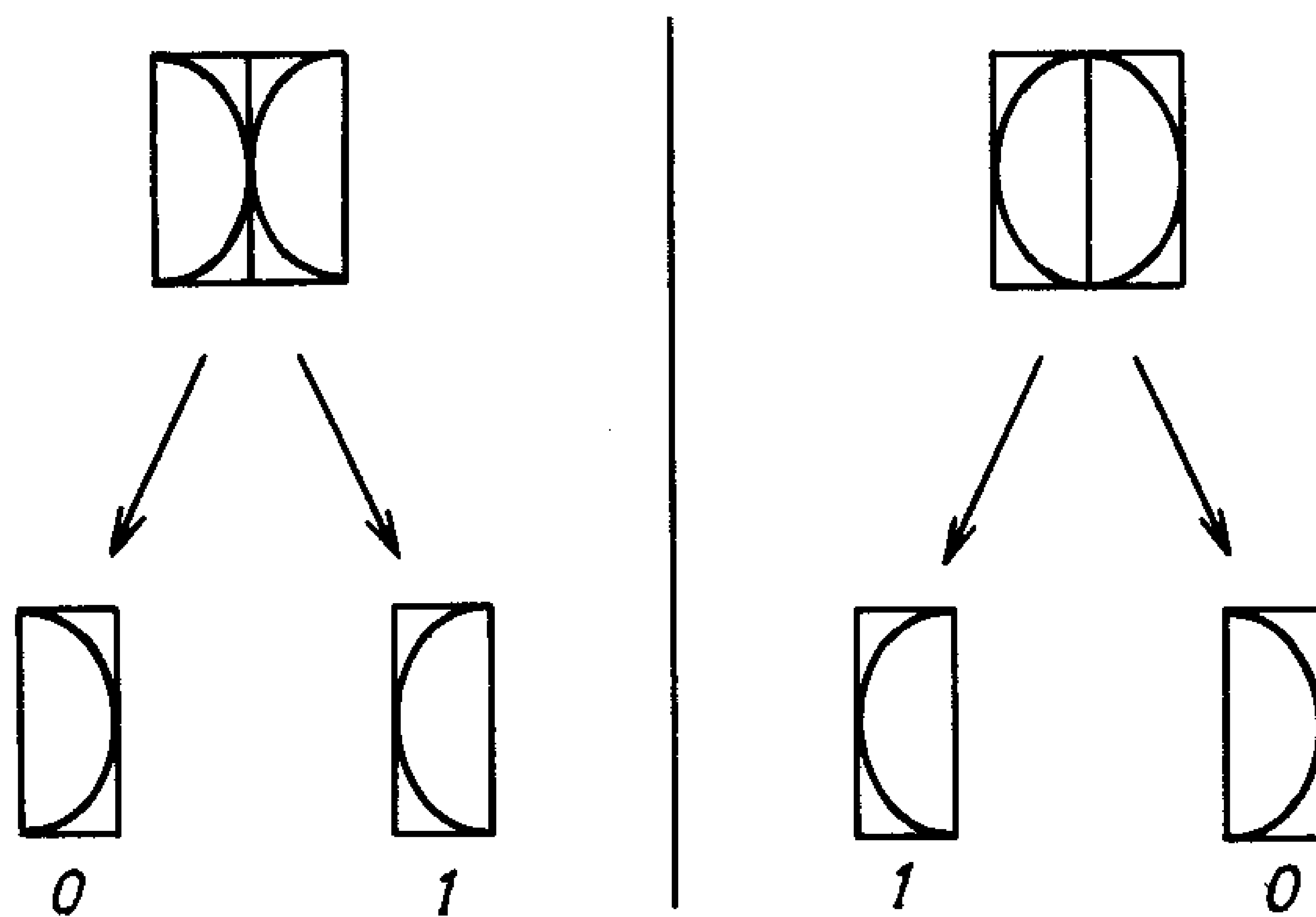


Рис. 12.1. Пример разложения букв X и O – выделения характерных особенностей, которым в соответствие ставятся числа 0 или 1.

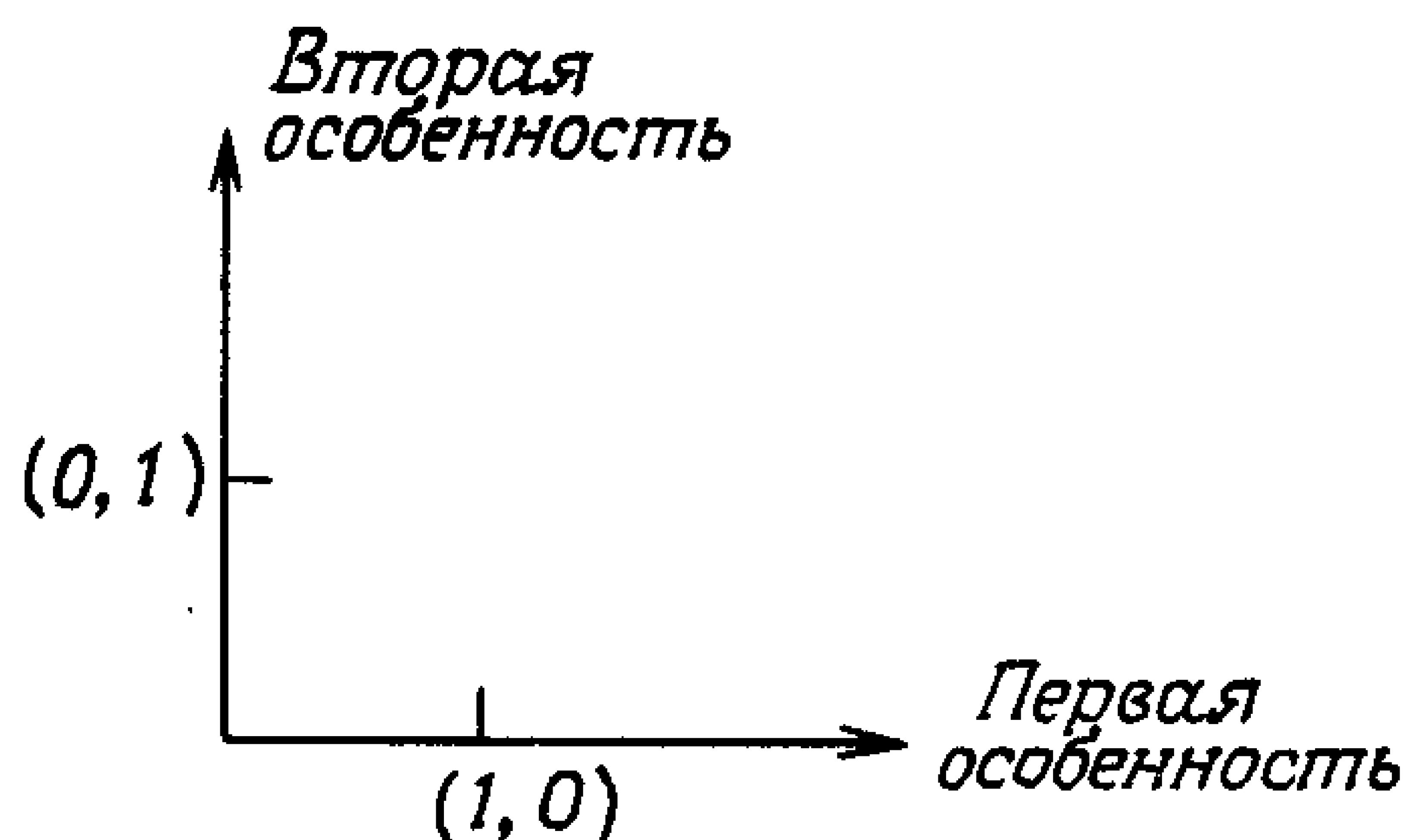


Рис. 12.2. Пространство особенностей для символов рис. 12.1. Здесь точка (1, 0) соответствует букве O, точка (0, 1) – букве X.

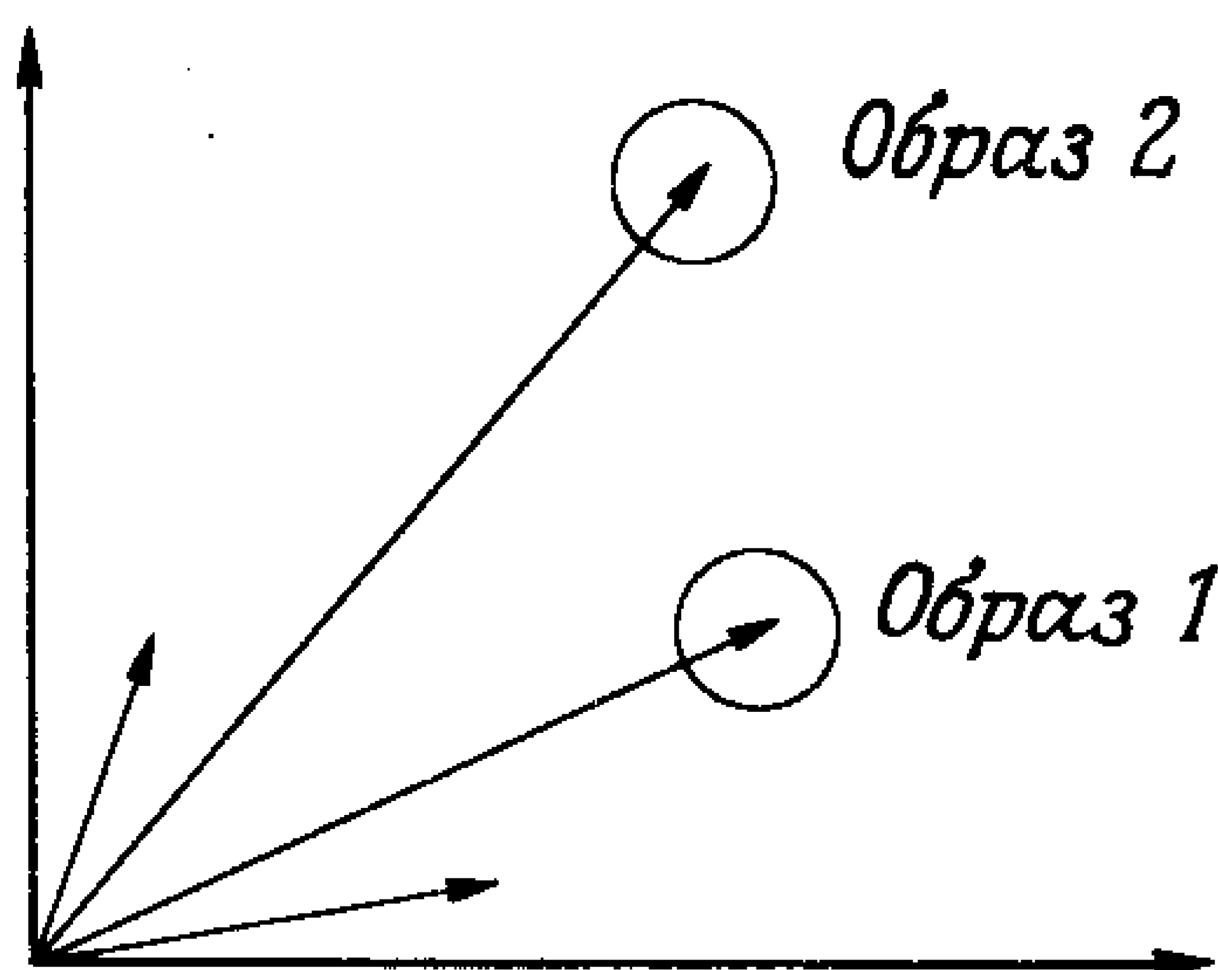


Рис. 12.3. Визуализация пространства характерных особенностей. Концы векторов указывают образ-прототип, в то время как точки, расположенные в окрестности концов, соответствуют образам, слегка отличающимся от прототипа и все еще сохраняющим сходство с ним.

ние образов происходит в два этапа. Выбирая признаки, мы должны ставить им в соответствие вектор $\mathbf{q}$, компоненты которого мы определяем, производя соответствующие измерения. Затем требуется установить в пространстве признаков, какой структуре отвечает построенный нами вектор $\mathbf{q}$.

К сожалению, при практическом осуществлении этой программы возникают трудности. Начать с того, что признаки могут быть выбраны различными способами. В этой связи возникает проблема адекватного выбора признаков. Под адекватным мы понимаем такой выбор признаков, при котором пространство признаков имеет не слишком высокую размерность, хотя все структуры при этом отчетливо различимы. Невысокая размерность пространства означает, что обработке подлежит не слишком большое количество информации. Кроме того, некоторые трудности возникают и в связи с самими измерениями (или структурами), поскольку измерения могут вносить случайный шум (а структуры – обладать внутренними флуктуациями). Это означает, что мы имеем дело не с детерминистическим вектором $\mathbf{q}$, а с вектором, компонентами которого являются случайные величины с некоторой функцией распределения вероятности.

В дальнейшем нас будут интересовать две задачи.

1) Как можно, исходя из экспериментальных данных, определить эту функцию распределения вероятности?

2) Как можно выделить признаки образа, чтобы подлежащая обработке информация была минимальна?

Чтобы получить функцию распределения вероятности $P(\mathbf{q})$, обратимся к принципу максимума информации. Предположим для этого, что мы произвели измерения над корреляционными функциями:

$$\langle q_i q_j \rangle = Q_{ij} \qquad (12.1)$$

по ансамблю структур, которые требуется классифицировать или исследовать. Не уменьшая общности, можно предположить, что среднее $\langle \mathbf{q} \rangle$ равно нулю, т. е.

$$\langle \mathbf{q} \rangle = 0. \qquad (12.2)$$

Если это не так, то мы всегда можем ввести новые векторы структур, просто вычитая ненулевые средние из исходных векторов. Применяя принцип максимума информационной энтропии, получаем

$$P(\mathbf{q}) = N \exp(-\mathbf{q}^T M \mathbf{q}), \qquad (12.3)$$

где в качестве ограничений мы использовали величины (12.1) и (12.2). Из множителей Лагранжа можно составить матрицу M, связанную с корреляционной матрицей Q соотношением

$$M = Q^{-1}, \qquad (12.4)$$

где матрица Q определяется выражением (12.1).

Преобразуем функцию (12.3) так же, как в разд. 6.3 при выводе структуры, связанной с модами вблизи неравновесного фазового перехода. Прежде всего найдем собственные векторы $\mathbf{u}_j$ и собственные значения λ_j уравнения

$$Q\mathbf{u}_j = \lambda_j \mathbf{u}_j. \tag{12.5}$$

Эти собственные векторы диагонализуют матрицу Q. Покажем, что эти же векторы диагонализуют также и матрицу M. Образуем для этого матрицу U из собственных векторов $\mathbf{u}_j$:

$$(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) = U, \tag{12.6}$$

где элементы матрицы U определяются соотношением

$$u_{kj} = U_{kj}. \tag{12.7}$$

Индекс k позволяет различать отдельные компоненты вектора, а индекс j – различные собственные векторы. Матрица U позволяет записать отдельно уравнения (12.5) в виде единого матричного уравнения

$$QU = U\Lambda, \tag{12.8}$$

где матрица Λ имеет элементы

$$\Lambda_{kj} = \lambda_j \delta_{jk}. \tag{12.9}$$

Матричное уравнение (12.8) в покомпонентной записи имеет вид

$$\sum_k Q_{ik} U_{kj} = \sum_l u_{il} \Lambda_{lj} = \lambda_j u_{ij}. \tag{12.10}$$

В дальнейшем мы будем предполагать, что матрица Q невырожденна (по существу это предположение содержится в соотношении (12.4)). Умножая матричное уравнение (12.8) слева на Q^{-1}, получаем

$$U = Q^{-1} U\Lambda. \tag{12.11}$$

При умножении справа на Λ^{-1} оно переходит в уравнение

$$U\Lambda^{-1} = Q^{-1} U. \tag{12.12}$$

Соотношение (12.4) позволяет записать уравнение (12.12) в виде

$$MU = U\tilde{\Lambda}, \tag{12.13}$$

где для краткости мы ввели обозначение

$$\tilde{\Lambda} = \Lambda^{-1}. \tag{12.14}$$

Так как Q – симметричная матрица, справедливо соотношение

$$U^T U = 1, \tag{12.15}$$

где U^T – матрица, транспонированная относительно матрицы U.

Разложим произвольный вектор $\mathbf{q}$ по собственным векторам $\mathbf{q}_j$. В общем случае это разложение можно записать в виде

$$\mathbf{q} = U\xi, \tag{12.16}$$

или, более подробно, как

$$\mathbf{q} = \sum_j \xi_j \mathbf{u}_j. \tag{12.17}$$

Подчеркнем, что вектор $\mathbf{q}$ – случайная величина, тогда как собственные векторы $\mathbf{u}_j$ – величины детерминистические. Следовательно, входящий в разложения (12.16) и (12.17) вектор ξ должен иметь компоненты, которые являются случайными величинами. Мы хотим сейчас выразить показатель экспоненты в правой части формулы (12.3), т. е. величину

$$\mathbf{q}^T M \mathbf{q}, \tag{12.18}$$

через новую переменную ξ. Для этого транспонируем вектор (12.17), т. е. возьмем вектор

$$\mathbf{q}^T = \xi^T U^T. \tag{12.19}$$

Тогда

$$\mathbf{q}^T M \mathbf{q} = \xi^T U^T M U \xi. \tag{12.20}$$

Используя уравнения (12.13) и (12.14), запишем соотношение (12.20) в виде

$$MU = \xi^T U^T U \Lambda^{-1} \xi, \tag{12.21}$$

который с учетом соотношений (12.15) и (12.14) допускает дальнейшее упрощение:

$$\mathbf{q}^T M \mathbf{q} = \xi^T \tilde{\Lambda} \xi. \tag{12.22}$$

Как хорошо известно из теории преобразований функций распределения вероятности, мы должны преобразовывать не только показатель экспоненты функции распределения вероятности, но и элемент объема. Это преобразование определяется якобианом, который в нашем случае оказывается равным

$$\operatorname{Det} U = 1, \tag{12.23}$$

поэтому преобразование оставляет элементы объема неизменными. В результате мы получаем новую функцию распределения вероятности: вместо формулы (12.3) она определяется выражением

$$\tilde{P}(\xi) = \mathcal{N} \exp(-\xi^T \tilde{\Lambda} \xi). \tag{12.24}$$

Учитывая, что $\tilde{\Lambda}$ – диагональная матрица, запишем функцию распределения в виде

$$\tilde{P}(\xi) = \mathcal{N} \exp(-\sum_j \tilde{\lambda}_j \xi_j^2). \tag{12.25}$$

Правую часть можно разложить в произведение

$$\tilde{P}(\xi) = \prod_j P_j(\xi_j), \tag{12.26}$$

где каждый множитель определяется выражением

$$P_j(\xi_j) = \mathcal{N} \exp(-\tilde{\lambda}_j \xi_j^2) \tag{12.27}$$

и нормировочный коэффициент $\mathcal{N}_j$ выбирается так, чтобы множитель P_j был нормирован на единицу. Функция распределения, представимая в виде распределения (12.26), где множители определяются выражениями (12.27), порождает корреляционную функцию

$$\langle \xi_j \xi_k \rangle = \lambda_j \delta_{jk}, \tag{12.28}$$

т. е. корреляционная матрица становится теперь диагональной. Заметим также, что наименьшее собственное число $\tilde{\lambda}_j$ соответствует наибольшему λ_j.

Выясним теперь, чего мы достигли преобразованием от $\mathbf{q}$ к ξ с точки зрения информации. В совершенно общем случае информация определяется выражением

$$i = -\int d^N q P(\mathbf{q}) \ln P(\mathbf{q}), \tag{12.29}$$

которое с помощью приведенной выше процедуры может быть приведено к виду

$$i = -\int d^N \xi \, \tilde{P}(\xi) \ln \tilde{P}(\xi). \tag{12.30}$$

Соотношение (12.25) позволяет записать выражение (12.30) как сумму отдельных вкладов:

$$i = -\sum_j \int d\xi_j P_j(\xi_j) \ln P_j(\xi_j), \tag{12.31}$$

или кратко:

$$i = \sum_j i_j. \tag{12.32}$$

Формула (12.27) позволяет сразу же вычислить информацию, выразив ее через собственные значения $\tilde{\lambda}_j$:

$$i_j = -\frac{1}{2}\ln \tilde{\lambda}_j + \frac{1}{2}\ln \pi + \frac{1}{2}. \tag{12.33}$$

Так как $\tilde{\lambda}_j = 1/\lambda_j$, выражение (12.33) можно представить также в следующем виде:

$$i_j = \frac{1}{2}\ln\lambda_j + \frac{1}{2}\ln\pi + \frac{1}{2}. \qquad (12.34)$$

Из соотношений (12.33) и (12.34) мы заключаем, что наибольший вклад в информацию дает наибольшее собственное значение λ_j и что полная информация представима в виде суммы некоррелированных вкладов парциальных информаций. Обсудим теперь, что это означает для распознавания образов.

Мы начали с отобранных нами признаков, которым поставили в соответствие вектор **q**. Выбирая различные линейные комбинации вида (12.27), мы производим выбор различных признаков. Иначе говоря, мы извлекаем признаки нового типа. Новые признаки обладают тем весьма приятным свойством, что информация становится суммой парциальных вкладов (12.34). На этой стадии мы уже можем обратиться к методам, развитым в теории численного распознавания образов. Разложение (12.17) с векторами **u**, выбранными, как в характеристическом уравнении (12.5), называется разложением Карунена–Лоэва. В этом контексте величина i называется энтропией населенности. На языке теории распознавания образов наша теория допускает следующую интерпретацию: разложение Карунена–Лоэва дает нам метод систематического разложения образа на некоррелированные компоненты (признаки). Как будет показано ниже, это позволяет уменьшить число компонент, или признаков. Соответствующая разложению Карунена–Лоэва (К-Л) система координат $\mathbf{u}_j$ детерминистическая. Компоненты, или, точнее говоря, коэффициенты, разложения К-Л содержат всю информацию, необходимую для восстановления образа. Коэффициенты с наибольшим собственным значением λ_j содержат наибольшую информацию относительно вектора образа **q**.

Таким образом, для распознавания образа нам необходимо учитывать только коэффициенты с наибольшими собственными значениями, в то время как коэффициенты с малой дисперсией (что эквивалентно малому собственному значению) содержат пренебрежимо малое количество информации. Иначе говоря, можно утверждать, что преобразование вектора **q** в систему координат К-Л сопровождается сокращением информации. Коль скоро система координат К-Л определена, она одинакова для всех образов, подлежащих разбиению на классы (образы, принадлежащие одному и тому же классу, считаются эквивалентными, или неразличимыми).

Итак, носителями всей информации, позволяющей дискриминировать образы, являются коэффициенты ξ_j разложения, причем наибольшее количество информации содержится в небольшом числе этих коэффициентов ξ. Вычислительные трудности, связанные с решением нашей задачи, удается уменьшить, если последующее реше-

ние принимать в пространстве признаков меньшей размерности. Использование информации обладает еще одной очень приятной особенностью: чем больше дисперсия собственных значений λ_j, тем меньше энтропия соответствующей популяции i, тем больше сокращение информации. Наоборот, при равномерно распределенных по величине собственных значениях λ_j мы имеем сильную неопределенность, большую энтропию, и наблюдение одного коэффициента ξ лишь незначительно уменьшает неопределенность или энтропию.

Описанная процедура допускает интересную интерпретацию с точки зрения неравновесных фазовых переходов. Если система находится сначала в неупорядоченном состоянии и приближается к точке перехода, то флуктуации параметров порядка нарастают и становятся особенно большими по сравнению со всеми другими модами. Именно при такой интерпретации можно утверждать, что коэффициенты ξ_j с наибольшей дисперсией или с наибольшими собственными значениями λ_j соответствуют параметрам порядка, определяющим образ. В этой области функция распределения вероятности вполне характеризуется своими моментами первого и второго порядка в полном согласии с нашей теорией, изложенной в этом разделе. Следующий раздел мы посвятим распознаванию образов «выше точки перехода» (что позволит нам воспользоваться интерпретацией распознавания образов как фазового перехода).

12.2. Алгоритм распознавания образов

Как уже упоминалось в разд. 12.1, мы описываем образ N-мерным вектором $\mathbf{q}$, отдельные компоненты которого являются признаками, закодированными с помощью чисел. Мы предполагаем также, что число M прототипов образов задано и может быть представлено векторами состояний

$$\mathbf{v}^{(k)} = \begin{pmatrix} v_1^{(k)} \\ \cdot \\ \cdot \\ \cdot \\ v_N^{(k)} \end{pmatrix}, \tag{12.35}$$

где индекс k принимает значения от 1 до M. В зависимости от числа векторов $\mathbf{v}^{(k)}$ их линейная оболочка совпадает либо со всем пространством признаков, либо с его подпространством. В последнем случае приходится вводить дополнительные («лишние») векторы. Образцы (12.35) хранятся в компьютере в таком виде, как это будет описано ниже. Когда предъявляется образ, который должен быть распознан компьютером, из-за шума, неполноты данных (отсутствия некоторых признаков) и т. д. этот образ может не совпадать ни с одним из векторов (12.35), но оказаться достаточно близко к одному

из прототипов образов, т. е. к одной из осей пространства признаков. Наша основная идея сводится к следующему. Условимся описывать предъявленный образ вектором $\mathbf{q}(0)$, где аргумент 0 относится к начальному моменту времени $t = 0$ процесса распознавания образов. Затем нам необходимо найти уравнения для зависящего от времени вектора $\mathbf{q}(t)$, которые описывали бы эволюцию этого вектора из начального состояния $\mathbf{q}(0)$ в конечное состояние $\mathbf{q}_f$, согласующееся с одним из прототипов векторов $\mathbf{v}^{(k)}$. Эти уравнения будут построены нами с таким расчетом, чтобы вектор $\mathbf{v}^{(k)}$, о котором идет речь, был тем самым вектором, к которому вектор $\mathbf{q}(0)$ подходит ближе всего, т. е. величина $(\mathbf{v}^{(k)+} \cdot \mathbf{q}(0))/(|\mathbf{v}^{(k)}|\,|\mathbf{q}(0)|)$ (12.35) имела наименьшее значение. Мы предполагаем, что

$$(\mathbf{v}^{(k)+} \cdot \mathbf{v}^{(k')}) = \delta_{kk'}.$$

Предполагаемая процедура приводит к двумя приятным следствиям. (1) Даже если начальный образ неполон, он может быть дополнен, т. е. наша система действует как «ассоциативная память». Например, по телефонному справочнику фамилия абонента (соответствующая $\mathbf{q}(0)$) дополняется номером телефона, и пара «фамилия + номер телефона» соответствует вполне определенному вектору $\mathbf{v}^{(k)}$. (2) Если мы храним в качестве прототипов образов $\mathbf{v}^{(k)}$ пары «портрет + фамилия» и предъявляем только портрет ($\mathbf{q}(0)$), то система превращает $\mathbf{q}(0)$ в $\mathbf{q}(t)$ – «портрет + фамилия». Если же мы узнаем фамилию того лица, чей портрет предъявлен, то этот портрет можно считать опознанным (машиной).

Наши уравнения, с помощью которых осуществляется все это, имеют следующий вид:

$$\dot{\mathbf{q}} = -\mathrm{grad}_{\mathbf{q}} V + F(t). \quad (12.36)$$

Входящий в них потенциал V определяется уравнением

$$V(\mathbf{q}) = V_0 + V_1 + V_2, \quad (12.37)$$

где

$$V_0 = -\frac{1}{2}\mathbf{q}^+ \sum_k \lambda \mathbf{v}^{(k)} \cdot \mathbf{v}^{(k)+}\, \mathbf{q}, \quad (12.38)$$

$$V_1 = \sum_{k \neq k'} C_{kk'} (\mathbf{v}^{(k)+}\, \mathbf{q})^2 (v^{(k')+}\, \mathbf{q})^2, \quad (12.39)$$

$$C_{kk'} > 0 \quad (12.40)$$

и

$$V_2 = C\mathbf{q}^4. \quad (12.41)$$

Слагаемое V_0 «затягивает» вектор $\mathbf{q}$ в подпространство, натянутое на прототипы образов. На языке теории неравновесных фазовых пере-

ходов это просто *пространство параметров порядка.* Слагаемое V_1 служит для дискриминации векторов $\mathbf{q}$ в подпространстве прототипов образов. В этом можно убедиться, если воспользоваться тем свойством, что минимум функции V_1 выбирается в соответствии с условием

$$V_{1,\min} = 0 \qquad \text{при} \qquad \mathbf{q} \,||\, \mathbf{v}^{(k_0)}, \tag{12.42}$$

т. е. когда вектор состояния $\mathbf{q}$ становится параллельным одному из прототипов векторов и таким образом начальный вектор состояния $\mathbf{q}(0)$ отождествляется с соответствующим прототипом образа. Наконец, слагаемое V_2 обеспечивает насыщение, т. е. величина $|\mathbf{q}|$ в конце концов притягивается к аттрактору – неподвижной точке на оси $\mathbf{v}^{(k_0)}$. Постоянная λ в соотношении (12.38) играет роль управляющего параметра: $\lambda < 0$ определяет область ниже «порога», $\lambda > 0$ – область выше порога. В этом разделе мы будем предполагать, что $\lambda > 0$.

Итак, вектор $\mathbf{q}$ необходимо привести к совпадению с одним из прототипов векторов. Чтобы упростить эту операцию, мы добавили в уравнения (12.36) флуктуирующую силу $\mathbf{F}$. Как обычно, мы предполагаем, что флуктуирующие силы обладают следующими свойствами:

$$\langle \mathbf{F}(t) \rangle = 0, \tag{12.43}$$

$$\langle F_i(t) F_j(t') \rangle = Q \delta_{ij} \delta(t - t'). \tag{12.44}$$

Наш подход к распознаванию образов был проверен с помощью численного эксперимента. Были изготовлены фотографии нескольких человек. Их фамилии были закодированы с помощью букв в правом верхнем углу каждой фотографии. Затем фотография переводилась в цифровую форму, для чего снимок разбивался на квадратную сетку размером 60×60 ячеек. Каждая ячейка соответствует компоненте вектора $\mathbf{v}^{(k)}$. На рис. 12.4 показано несколько типичных примеров. Затем компьютеру, который решал уравнения (12.36), предъявлялся образ-тест $\mathbf{q}$ либо в виде отдельной буквы, либо в виде части лица. Результаты распознавания представлены на рис. 12.5. Переходя от векторов $\mathbf{v}^{(k)}$ и $\mathbf{q}$ к их фурье-образам и используя абсолютные величины фурье-образов в качестве новых переменных, весь процесс распознавания образов можно сделать инвариантным относительно трансляций. Затем мы предъявляли компьютеру различные изображения (рис. 12.6). Первой компьютер узнал женщину. Затем он узнал частично закрытое мужское лицо. (Заметим, что изображенных на групповой фотографии 5 человек компьютер также узнал постепенно, шаг за шагом.)

В дальнейшем нас будет интересовать и другой аспект распознавания образов, а именно в приведенные выше уравнения нам приходилось вводить прототипы векторов, теперь же мы хотим показать, каким путем компьютер можно научить распознавать образы. В качестве «наглядного пособия» мы используем аналогию между

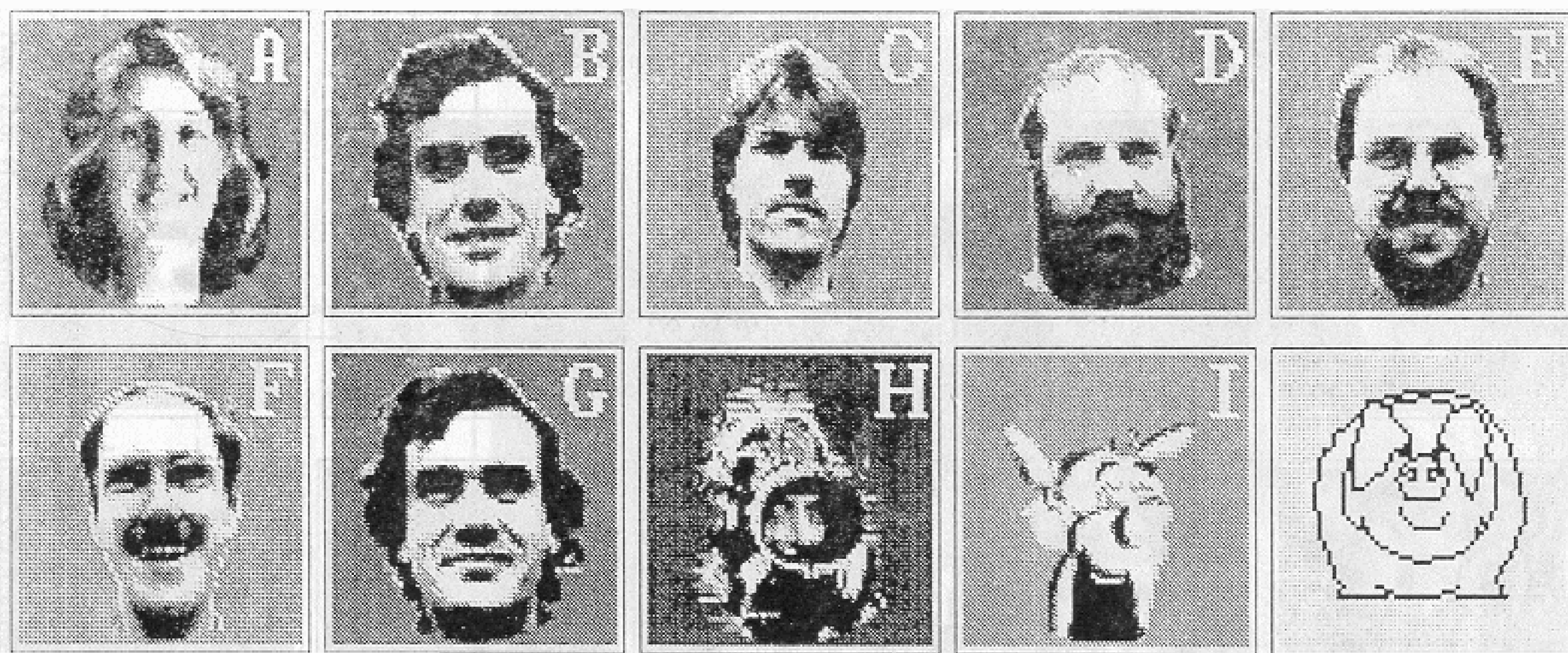


Рис. 12.4. Прототипы образов, хранящиеся в памяти компьютера.

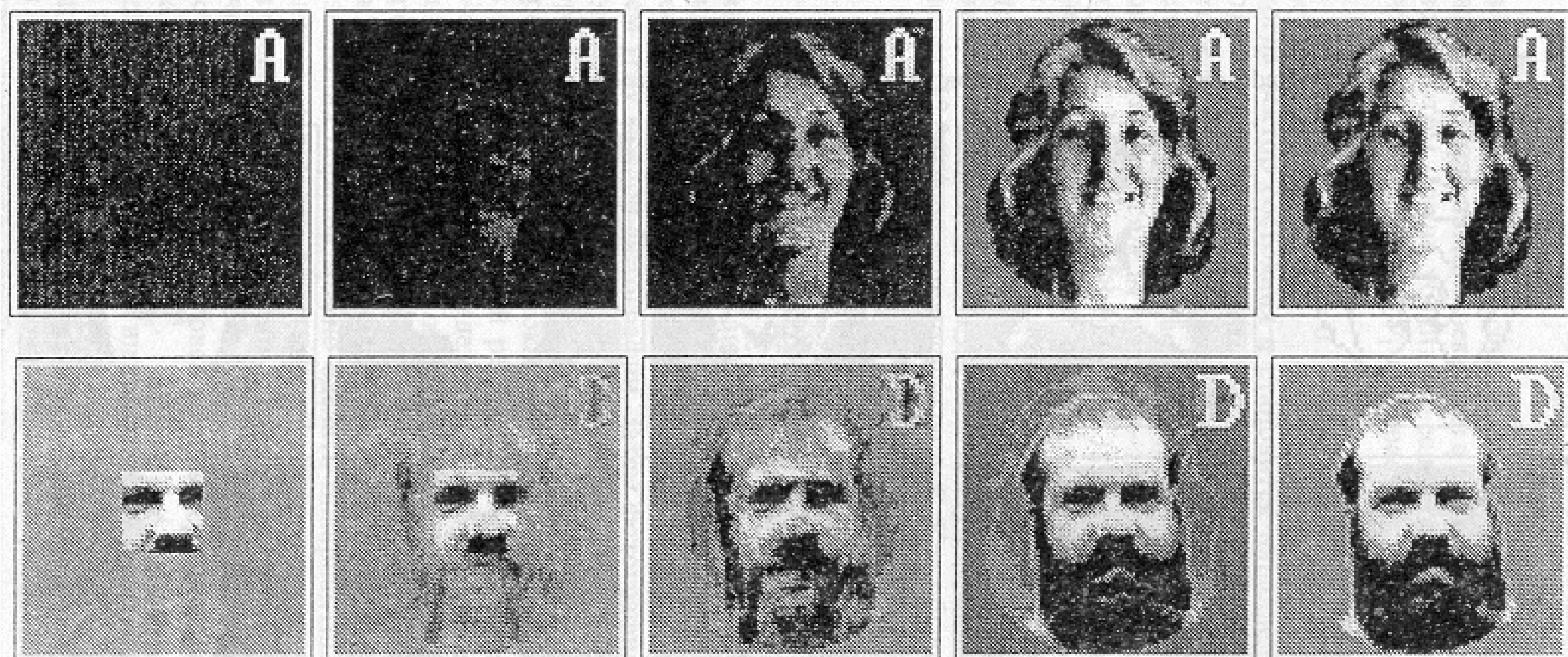


Рис. 12.5. Восстановление изображения.



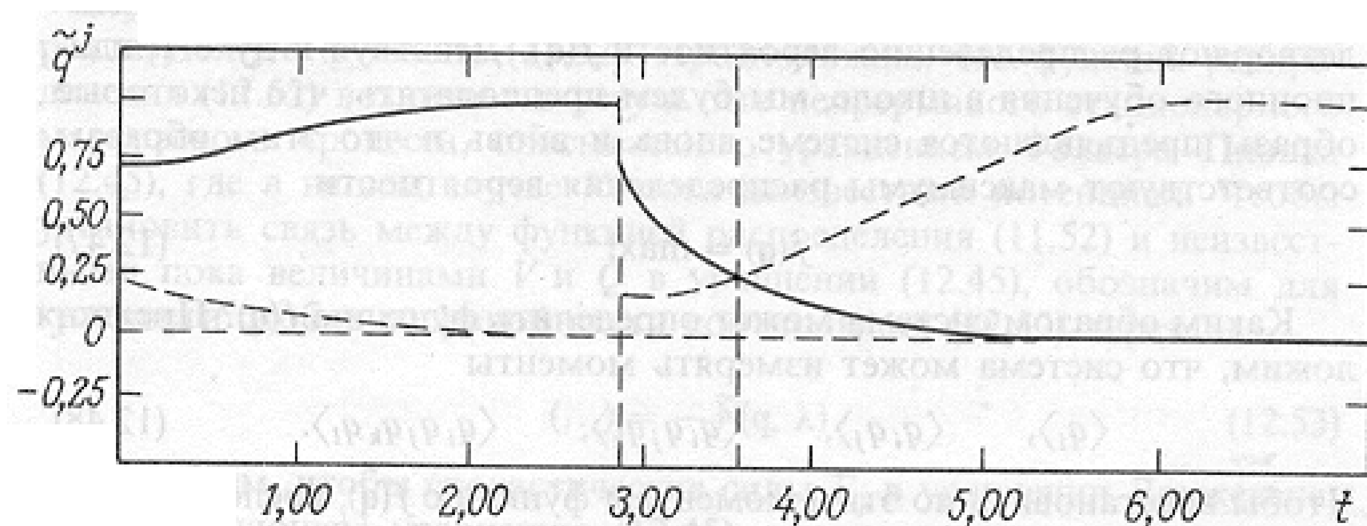


Рис. 12.6. Изображение, распознанное компьютером.

распознаванием образов, описываемым уравнениями (12.36), и неравновесными фазовыми переходами.

Начнем с того, что преобразуем уравнения (12.36) типа уравнения Ланжевена в уравнение Фоккера–Планка:

$$\dot{f} = \sum_k \frac{\partial}{\partial q_k}\left(\frac{\partial V}{\partial q_k}\cdot f\right) + \frac{Q}{2}\sum_k \frac{\partial^2 f}{\partial q_k^2}. \tag{12.45}$$

Стационарное решение его находится без труда и имеет вид (см. (2.24))

$$f = N\exp\left(-\frac{2V}{Q}\right). \tag{12.46}$$

Оно означает, что стационарное явление полностью определяется потенциалом V (представимым в виде суммы трех членов (12.37)) и величиной флуктуаций Q. Максимумы решения $f(\mathbf{q})$ соответствуют минимумам потенциала $V(\mathbf{q})$. Как показывает простой анализ, минимумы потенциала $V(\mathbf{q})$ расположены в точках $\mathbf{q} = \mathbf{v}^{(k)}$, т. е совпадают с прототипами образов. Сказанное справедливо при $\lambda > 0$. Что же касается области $\lambda < 0$, то в ней минимум потенциала расположен в точке $\mathbf{q} = \mathbf{0}$, т. е. даже если системе будут предъявлены первоначальные образы, они все равно исчезнут.

12.3. Основные принципы конструирования синергетического компьютера

Предположим, что система может получать некоторое множество данных, описываемых вектором состояния $\mathbf{q}$. Предположим также, что детектирующая система состоит из элементов j и что элемент j измеряет компоненту q_j (признак) вектора $\mathbf{q}$. В дальнейшем мы выясним, какими свойствами должны обладать эти элементы *и связи между ними* для того, чтобы они могли распознавать образы. Если существует стационарный процесс, то поступающие сигналы удовлетворяют распределению вероятности $f(\mathbf{q})$. Действуя в духе традиционного обучения в школе, мы будем предполагать, что некоторые образы предъявляются системе вновь и вновь и что этим образам соответствуют максимумы распределения вероятности

$$f(\mathbf{q}) = \max! \tag{12.47}$$

Каким образом система может определить функцию $f(\mathbf{q})$? Предположим, что система может измерять моменты

$$\langle q_i \rangle, \qquad \langle q_i q_j \rangle, \qquad \langle q_i q_j q_k \rangle, \qquad \langle q_i q_j q_k q_l \rangle. \tag{12.48}$$

Чтобы восстановить по этим моментам функцию $f(\mathbf{q})$, воспользуемся принципом максимума информационной энтропии и попытаемся понять, какой минимальный порядок моментов необходим для полу-

чения разумной функции f. Если ограничиться только линейными моментами, то, согласно принципу максимума информационной энтропии, функция f должна иметь следующий вид: $f(\mathbf{q}) = N\exp(-\lambda\mathbf{q})$, где λ - вещественное число. Совершенно ясно, что функция $f(\mathbf{q})$ не может быть нормирована в пространстве векторов $\mathbf{q}$, если переменные $\mathbf{q}$ принимают значения от минус до плюс бесконечности, поэтому ограничиться только линейными моментами невозможно. Если использовать моменты до второго порядка включительно, то общий вид функции распределения окажется следующим:

$$f(\mathbf{q}) = N\exp(\bar{\boldsymbol{\alpha}}\mathbf{q} + \bar{\mathbf{q}}\boldsymbol{\alpha} - \bar{\mathbf{q}}B\mathbf{q}). \tag{12.49}$$

Вводя новую переменную ξ с помощью соотношения

$$\mathbf{q} = \xi + B^{-1}\boldsymbol{\alpha}, \tag{12.50}$$

мы можем записать функцию (12.49) несколько иначе:

$$\tilde{f}(\xi) = N'\exp(-\bar{\xi}B\xi). \tag{12.51}$$

Так как B - положительно-определенная матрица, максимум функции (12.51) может быть достигнут только при одном значении переменной ξ, а именно при $\xi = 0$.

Производя замену переменных (12.50), мы сознаем, что в нашей модели существует только один максимум, т. е. хранится только один образ. В действительности для распознавания обычно предъявляется *набор образов*. Это противоречие приводит нас к необходимости рассматривать моменты до четвертого порядка включительно. Для простоты будем считать, что моменты нечетного порядка равны нулю. Функция распределения при этом в соответствии с принципом максимума информационной энтропии имеет следующий вид:

$$f(\mathbf{q}) = N\exp\Big(-\sum_{i,j}\lambda_{ij}q_iq_j - \sum_{i,j,m,n}\lambda_{ijmn}q_iq_jq_mq_n\Big). \tag{12.52}$$

Теперь мы хотим показать, каким образом можно построить сеть, реализующую функцию (12.52). Предположим, что функция распределения (12.52) возникает в результате непрерывного стационарного марковского процесса, описываемого уравнением Фоккера-Планка (12.45), где в нашем случае V - пока неизвестный потенциал. Чтобы установить связь между функцией распределения (11.52) и неизвестными пока величинами V и Q в уравнении (12.45), обозначим для краткости скобки в правой части соотношения (12.52) через

$$(\ldots) = -\tilde{V}(\mathbf{q}, \lambda) \tag{12.53}$$

и потребуем, чтобы стохастические силы F_j в уравнении Ланжевена, соответствующем уравнению (12.45), удовлетворяли соотношениям (12.43), (12.44). Пусть, кроме того,

$$\tilde{V} = \frac{2V}{Q}. \tag{12.54}$$

Уравнение Ланжевена, соответствующее уравнению Фоккера–Планка (12.45), имеет, вообще говоря, такой же вид, как уравнение (12.36), а именно:

$$\dot{\mathbf{q}}_j = -\frac{\partial V}{\partial q_j} + F_j(t). \tag{12.55}$$

Но потенциал V теперь определяется соотношениями (12.53), (12.54), т. е. результатами измерений. Мы получаем

$$-\frac{\partial V}{\partial q_i} = Q\left\{\sum_j \lambda_{ij} q_j + 2 \sum_{j,\,m,\,n} \lambda_{ijmn} q_i q_j q_m q_n\right\}. \tag{12.56}$$

Заметим, что уравнение (12.55), в котором первый член в правой части определяется соотношением (12.56), принимает особенно простой вид для сети, описание которой было приведено в предыдущем разделе. Иначе говоря, можно построить сеть, реализующую функцию распределения $f(\mathbf{q})$, если известны второй и четвертый моменты этой функции. Назначение сети состоит только в том, чтобы передавать значение величины q_j, умноженное на «синаптическую силу» λ_{ij}, от элемента i к элементу j [первый член в правой части соотношения (12.56)] или передавать значение $q_j q_m q_n$, умноженное на λ_{ijmn}, от элементов j, m, n к элементу i [второй член в правой части соотношения (12.56)], после чего в соответствующем элементе i производится суммирование.

Справедлив еще один совершенно общий и важный вывод. Если сеть может измерять только вполне определенный набор корреляционных функций (12.48) (или более высокого порядка), то в выражение (12.56) необходимо лишь включить соответствующие члены. Существование в потенциальной динамике состояний-аттракторов приводит к тому, что любое начальное состояние оказывается притянутым к одному из локальных минимумов потенциала $V(\mathbf{q})$, которые расположены в точках, соответствующих прототипам образов. Если начальное состояние достаточно близко к одному из таких минимумов, то этот минимум реализуется и первоначально неполный образ дополняется, т. е. весь механизм действует как своего рода ассоциативная память и, следовательно, обладает способностью распознавать образы.

12.4. Обучение с помощью приращения информации

В этом разделе мы покажем, как система может определить величины коэффициентов λ_{ij} и λ_{ijmn}. Пусть $f(\mathbf{q})$ – заданная функция распределения поступающих в систему образов, а $\tilde{f}(\mathbf{q})$ – функция распределения,

порождаемая системой. Введем в качестве меры расстояния между этими двумя функциями прирост информации (информацию Кульбака)

$$K = \int f \ln\left(\frac{f}{\tilde{f}}\right) d^N q \geqslant 0, \tag{12.57}$$

где нам необходимо учитывать два ограничения:

$$\int f d^N q = 1, \tag{12.58}$$

$$\int \tilde{f} d^N q = 1. \tag{12.59}$$

Так как функция f задана, а выражение (12.57) можно представить в виде

$$K = \int f \ln f d^N q - \int f \ln \tilde{f} d^N q, \tag{12.60}$$

достаточно найти максимум выражения

$$\int f \ln \tilde{f} d^N q = \max! \tag{12.61}$$

Для определенности предположим, что функция f имеет вид

$$\tilde{f} = \exp[-\tilde{\lambda} - \sum_j \tilde{\lambda}_j V_j(\mathbf{q})], \tag{12.62}$$

где потенциал V_j может содержать полиномы относительно $\mathbf{q}$, степень которых не превышает некоторого заданного нами предела; $\tilde{\lambda}_j$ – параметры, которые можно варьировать. Левую часть выражения (12.61) (умноженную на -1) можно представить в виде

$$W \equiv -\int f \ln \tilde{f} d^N q = \tilde{\lambda} + \int f \sum_j \tilde{\lambda}_j V_j(\mathbf{q}) d^N q. \tag{12.63}$$

Мы предполагаем, что $\tilde{\lambda}_j$ подчинены эволюционной стратегии, например, в сети могут происходить флуктуации связности. Наибольшим изяществом отличается эволюционная стратегия, использующая градиенты некоторого потенциала. Выберем задаваемую соотношением (12.63) функцию W в качестве потенциала, подлежащего минимизации, и предположим, что

$$\tilde{\lambda} = \ln \int \exp[-\sum_j \tilde{\lambda}_j V_j(\mathbf{q})] d^N q. \tag{12.64}$$

Градиентная стратегия сводится к представлению множителей Лагранжа $\tilde{\lambda}_j$ в виде

$$\dot{\tilde{\lambda}}_j = -\gamma \frac{\partial W}{\partial \tilde{\lambda}_j}. \tag{12.65}$$

Чтобы вычислить правую часть этого равенства, подставим выражение (12.64) в формулу (12.63) и образуем производную

$$\frac{\partial W}{\partial \lambda_j} = -\left[\left\{\int \exp\left\{-\sum_j \tilde{\lambda}_j V_j(\mathbf{q})\right\} d^N q\right\}^{-1} \cdot \int V_j(\mathbf{q}) \exp\left\{-\sum_j \tilde{\lambda}_j V_j(\mathbf{q})\right\} d^N q\right] + \\ + \int f(\mathbf{q}) V_j(\mathbf{q}) d^N q. \qquad (12.66)$$

Как показывают простые рассуждения, выражение в правой части соотношения (12.66), заключенное в прямоугольные скобки, может быть интерпретировано как среднее значение

$$[\quad] = \langle V_j(\mathbf{q}) \rangle_{\tilde{f}}, \qquad (12.67)$$

в то время как последний член в правой части соотношения (12.66) есть среднее значение

$$\langle V_j(\mathbf{q}) \rangle_f. \qquad (12.68)$$

Поэтому равенство (12.65) может быть записано очень кратко как

$$\dot{\tilde{\lambda}}_j = \gamma\left(\langle V_j \rangle_{\tilde{f}} - \langle V_j \rangle_f\right), \qquad (12.69)$$

где первый член в правой части есть среднее от V_j по функции распределения $\tilde{f}$, а второй член есть среднее от того же потенциала V_j, но по функции распределения, задаваемой внешним миром. Сравнивая эти два значения, система может надлежащим образом отрегулировать связи $\tilde{\lambda}$ между своими элементами и таким образом научиться выполнять свою задачу. Соотношение (12.69) положено в основу машины, построенной Сейновским и др. и получившей название машины Больцмана. Эти авторы использовали корреляционные функции $q_i q_k$, в которых величины q могут принимать только два значения ± 1, как в модели спиновых стекол.

12.5. Обучение распознаванию процессов и ассоциативное действие

В предыдущих разделах было показано, как физическая система может научиться распознавать образы (паттерны) в сигналах и воспроизводить их распределение вероятности. В этом разделе я хочу показать, как система может научиться воспроизводить или распознавать *процессы*. В предыдущих разделах мы не указывали путь, по которому система переходит в конечные состояния-аттракторы, соответствующие первоначально предъявленным образам. В этом разделе будет показано, что система может даже научиться воспроизводить тот или иной путь. Мы будем исходить из основного предположения о том, что процессы, которые должна научиться распознавать система, являются марковскими. Условимся считать, что интересующий нас процесс стационарный и непрерывный. Марковский процесс характеризуется условными вероятностями:

$$P(\mathbf{q}_{i+1} \mid \mathbf{q}_i), \qquad (12.70)$$

вероятностью найти систему в точке $\mathbf{q}_{i+1}$ в момент времени $t_{i+1} = t_i + \tau$ при условии, что в момент времени t_i система находилась в точке $\mathbf{q}_i$. Обозначим компоненты векторов состояния в моменты времени t_{i+1} и t_i через

$$q_l(i+\tau), \qquad q_k(i). \tag{12.71}$$

Введем следующие условные моменты (первый и второй):

$$f_{1,l} = \langle q_l(i+\tau)\rangle_{\mathbf{q}(i)}, \tag{12.72}$$

$$f_{2,l,k} = \langle q_l(i+\tau)\, q_k(i+\tau)\rangle_{\mathbf{q}(i)}. \tag{12.73}$$

Воспользуемся некоторыми результатами, полученными в гл. 9. Применяя принцип максимума информационной энтропии к условной вероятности, получаем

$$P(\mathbf{q}(i+\tau)\,|\,\mathbf{q}(i)) = \exp\Big[-\lambda - \sum_l \lambda_l q_l(i+\tau) - \sum_{k,l} \lambda_{kl}\, q_k(i+\tau)\, q_l(i+\tau)\Big], \tag{12.74}$$

где множители Лагранжа могут зависеть от $\mathbf{q}(i)$. [Обратите внимание на то, что знаки по сравнению с формулой (9.23) изменены.] Введем вектор-столбец

$$\boldsymbol{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_N \end{pmatrix} \tag{12.75}$$

и матрицу

$$\Delta = (\lambda_{kl}) \tag{12.76}$$

и определим новый вектор $\mathbf{h}$ с помощью соотношения

$$\mathbf{h} = \frac{1}{2}\Delta^{-1}\boldsymbol{\lambda}. \tag{12.77}$$

Определим также величины G_{kl} и K_k: пусть

$$\lambda_{kl} = \frac{1}{\tau} G_{kl}(\mathbf{q}(i)) \tag{12.78}$$

и при $\tau \to 0$

$$h_k = q_k(i) + \tau K_k(\mathbf{q}(i)). \tag{12.79}$$

Как показано в разд. 9.2, уравнение Фоккера–Планка имеет в этом случае следующий вид:

$$\dot{f} = Lf, \tag{12.80}$$

где L–линейный оператор, определяемый соотношением

$$Lf = -\nabla_{\mathbf{q}}(K(\mathbf{q})f) + \frac{1}{4}\sum_{k,l}\frac{\partial^2}{\partial q_k \partial q_l}\{(G^{-1}(\mathbf{q}))_{kl}f\}. \tag{12.81}$$

Это уравнение Фоккера–Планка связано с исчислением Ито, которое позволяет нам легко построить уравнение Ито–Ланжевена (см. далее). В отличие от результатов, приведенных в разд. 12.3 и 12.4, флуктуирующие силы (по крайней мере в общем случае) становятся зависящими от $\mathbf{q}$. В остальном аналогия с результатами, приведенными в разд. 12.3, остается в силе, и, следовательно, мы можем построить нашу физическую систему. Если мы имеем физическую систему, в которой установлено только конечное число связей между элементами, то далее можно поступить по аналогии с разд. 9.3. Поскольку в случае нескольких переменных $\mathbf{q}_i$ предлагаемая процедура нетривиальна, мы опишем ее. В ряде физических приложений могут быть известны не векторы $\mathbf{q}_i$, а корреляционные функции того типа, который будет описан ниже. Мы предполагаем, что измерения производятся в стационарных условиях. В этом случае мы можем представить совместную вероятность в виде произведения условной вероятности и стационарного распределения вероятности

$$P(\mathbf{q}_{i+\tau}, q_i) = P(\mathbf{q}_{i+\tau} | \mathbf{q}_i) P_{\text{стац}}(\mathbf{q}_i). \tag{12.82}$$

Чтобы определить моменты или другие корреляционные функции, введем вспомогательные функции $U_j(\mathbf{q}_i)$, которые могут иметь, например, следующий вид:

$$U_j(\mathbf{q}_i) = U_{j,i} = q_{1,i}^{\mu_1} q_{2,i}^{\mu_2} \dots q_{N,i}^{\mu_N}; \qquad \mu_1 + \mu_2 + \dots + \mu_N = R. \tag{12.83}$$

Следует подчеркнуть, что индекс i относится к моментам времени, а остальные индексы 1, 2, ...–к компонентам вектора. Введем следующие ограничения:

$$\langle U_{j,i}^{(1)} \rangle, \tag{12.84}$$

$$\langle q_{k,i+\tau} U_{m,i}^{(2)} \rangle, \qquad k = 1, \dots, N, \qquad m = 1, 2, \dots, \tag{12.85}$$

$$\langle q_{k,i+\tau} q_{l,i+\tau} U_{n,i}^{(3)} \rangle; \qquad k = 1, \dots, N, \qquad l = 1, \dots, N, \qquad n = 1, 2, \dots . \tag{12.86}$$

Учитывая наши предыдущие результаты, мы сразу же находим стационарное распределение вида

$$P_{\text{стац}}(\mathbf{q}_i) = \exp(-\lambda_{\text{стац}} - \sum_j \lambda_{j,\text{стац}} U_{j,i}^{(1)}). \tag{12.87}$$

Воспользуемся теперь обобщением принципа максимума информационной энтропии на случай пространственно-временной области,

т. е. подчиним этому принципу многовременную совместную вероятность. Для излагаемого нами сейчас подхода достаточно рассмотреть двухвременную совместную вероятность, которая (как следует из принципа максимума информационной энтропии) принимает следующий вид:

$$P(\mathbf{q}_{i+\tau}, \mathbf{q}_i) = \exp[-\lambda_0 + A(\mathbf{q}_i) + \mathbf{B}(\mathbf{q}_i)\mathbf{q}_{i+\tau} + \bar{\mathbf{q}}_{i+\tau} C(\mathbf{q}_i)\mathbf{q}_{i+\tau}], \tag{12.88}$$

где мы воспользовались следующими обозначениями:

$$A(\mathbf{q}_i) = -\sum_j \lambda_j^{(1)} U_j^{(1)}(\mathbf{q}_i), \tag{12.89}$$

$$B_k(\mathbf{q}_i) = -\sum_m \lambda_{km}^{(2)} U_m^{(2)}(\mathbf{q}_i), \tag{12.90}$$

$$C_{kl} = -\sum_n \lambda_{kln} U_n^{(3)}(\mathbf{q}_i) \tag{12.91}$$

и λ-множители Лагранжа.

Соотношения (12.82), (12.87) и (12.88) позволяют нам записать формулу (12.70) в виде

$$P(\mathbf{q}_{i+\tau} | \mathbf{q}_i) = \exp[-\lambda + \tilde{A}(\mathbf{q}_i) + \mathbf{B}(\mathbf{q}_i)\mathbf{q}_{i+\tau} + \bar{\mathbf{q}}_{i+\tau} C(\mathbf{q}_i)\mathbf{q}_{i+\tau}], \tag{12.92}$$

где мы использовали для краткости обозначения

$$\lambda = \lambda_0 - \lambda_{\text{стац}}, \tag{12.93}$$

$$\tilde{A} = A + \sum_j \lambda_{j,\,\text{стац}} U_j^{(1)}(\mathbf{q}_i). \tag{12.94}$$

Сравнивая соотношение (12.92) с формулой (12.74), мы видим, что можем отождествить прежние множители Лагранжа с новыми следующим образом:

$$\lambda(\mathbf{q}_i) \leftrightarrow \lambda - A(\mathbf{q}_i), \tag{12.95}$$

$$\lambda_l(\mathbf{q}_i) \leftrightarrow -B_l(\mathbf{q}_i), \tag{12.96}$$

$$\lambda_{kl}(\mathbf{q}_i) \leftrightarrow -C_{kl}(\mathbf{q}_i). \tag{12.97}$$

После такого отождествления мы можем снова вывести уравнение Ито–Фоккера–Планка.

Если воспользоваться ограничениями в виде функций (12.83), то дрейфовый коэффициент и коэффициент диффузии оказываются полиномами. Из-за формальной аналогии между двухвременными функциями распределения вероятности и стационарной функцией распределения мы можем распространить процедуру обучения, изложенную в разд. 12.4, на случай, интересующий нас сейчас.

Выведем теперь уравнение Ито [см. уравнение (2.12)], соответствующее уравнению Фоккера–Планка. Это уравнение Ито имеет следующий вид:

$$dq_l(t) = K_l(\mathbf{q}(t))\,dt + \sum_m g_{lm}(\mathbf{q}(t))\,dw_m(t), \tag{12.98}$$

где стохастический процесс определяется [см. соотношения (2.13), (2.14)] соотношениями

$$\langle dw_m \rangle = 0, \tag{12.99}$$

$$\langle dw_m(t)\,dw_l(t) \rangle = \delta_{lm}\,dt. \tag{12.100}$$

Коэффициенты диффузии $(G^{-1})_{kl}$ уравнения Фоккера–Планка (12.80), (12.81) связаны с функциями g_{lm} формулой [см. соотношение (2.27)]:

$$\frac{1}{2}(G^{-1})_{kl} = \sum_m g_{km}\,g_{lm}. \tag{12.101}$$

Введем матрицы G и g, определив их соотношениями

$$(G_{kl}) = G; \qquad (g_{km}) = g. \tag{12.102}$$

Тогда формула (12.101) примет следующий вид:

$$G^{-1} = g\tilde{g}, \tag{12.103}$$

где $\tilde{g}$ означает транспонированную матрицу g. Уравнение (12.101) решается без труда, если предположить, что g – квадратная матрица и что матрицы g и G^{-1} симметричны. При этом уравнение (12.103) принимает вид

$$G^{-1} = g^2. \tag{12.104}$$

Матрицу g, удовлетворяющую уравнению (12.104), мы определим, если предположим, что матрицу G, которая положительно определена, можно диагонализовать с помощью матрицы U. Введем диагональную матрицу D с элементами D_l. Тогда

$$D = UG^{-1}\overline{U} = (Ug\overline{U})^2, \tag{12.105}$$

где

$$U\overline{U} = 1. \tag{12.106}$$

Отсюда немедленно следует, что

$$(Ug\overline{U})_{kl} = \delta_{kl}\sqrt{D_l}, \tag{12.107}$$

после чего найти матрицу g уже не составляет никакого труда.

Таким образом, мы показали, как может быть восстановлена физическая система, воспроизводящая марковский процесс. Следует ожидать, что такая система способна не только воспроизводить марковский процесс при полностью определенных начальных условиях, но и действовать как ассоциативная память, т. е. восстанавливать частично или полностью искаженные данные до полного множества правильных данных. Можно ожидать, что при этом эволюция такой системы будет происходить по траектории с исправлением ошибок. Такое поведение может быть названо ассоциативным действием.

Глава 13
Квантовые системы

13.1. Почему возникает необходимость в квантовой теории информации?

До сих пор мы занимались изучением классических систем, имеющих макроскопические размеры. В этой главе мы хотим рассмотреть системы, для описания которых требуется квантовая теория. Приведем несколько примеров таких систем, в которых существенную роль играют квантовые эффекты. Прежде всего здесь следует назвать лазер. Описание поведения отдельных атомов его активной среды требует квантовомеханических идей и методов, поле лазерного излучения порождается квантовомеханическими событиями – спонтанным и индуцированным испусканием света. В предыдущих разделах основной упор делался на так называемую полуклассическую теорию атома, которая может быть выведена из полностью квантовой теории с помощью принципа соответствия, подробное рассмотрение которого увело бы нас далеко в сторону.

Еще одним важным примером систем с квантовыми эффектами могут служить детали компьютера, непрерывная миниатюризация приводит к тому, что квантовые эффекты начинают играть заметную роль. Например, возникает необходимость в рассмотрении запирания одного или нескольких электронов в так называемых квантовых ямах. В диоде Ганна могут возникать осцилляции электронных токов, в полупроводниках – образовываться слоистые структуры. Электроны или дырки в различных запертых состояниях могут хранить информацию, и электроны могут образовывать пространственные или временные структуры. Биологические молекулы также могут быть хранителями информации, и со временем можно ожидать появления устройств, в которых для хранения и обработки информации молекулы будут использоваться наряду с полупроводниками. В этой связи было бы уместно попытаться распространить наши результаты на мир квантовомеханических процессов в таких микроскопических устройствах.

Как будет показано ниже, такая задача во многом реализуема. Вообще говоря, мы будем предполагать, что читатель знаком с основными понятиями квантовой теории, но для удобства напомним некоторые ее характерные особенности. Чтобы перейти от классической физики к квантовой механике, наблюдаемые q_i, например пространственные координаты и импульс частицы, необходимо заменить соответствующими операторами, которые в дальнейшем мы

также будем обозначать через q_i (табл. 13.1). Перевод на язык квантовой механики допускает не только классическая механика, но и теория электромагнитного поля. При квантовании электромагнитного поля амплитуда $E(x)$ электрического поля в точке с пространственной координатой x подлежит замене некоторым оператором $E(x)$. Если волновую функцию электрона $\psi(x)$, удовлетворяющую уравнению Шрёдингера, рассматривать как классическое поле, нам необходимо заменить ее некоторым оператором $\psi(x)$. Этот процесс известен под названием «вторичного квантования». Аналогичным образом становится оператором и плотность электронов (табл. 13.1).

Сделаем теперь еще один важный шаг вперед и включим в рассмотрение статистическую механику. В гл. 2 нам уже приходилось иметь дело с функцией распределения $f(\mathbf{q})$, где $\mathbf{q}$ – вектор динамических переменных, например пространственных координат частиц. В квантовой теории эту функцию распределения необходимо заменить матрицей плотности ρ. Если функция распределения удовлетворяет уравнению Фоккера–Планка, то матрица плотности удовлетворяет так называемому основному кинетическому уравнению. Ниже мы намереваемся предложить макроскопический вывод явного вида матрицы плотности ρ, полностью аналогичный приведенному выше «макроскопическому» выводу функции распределения $f(\mathbf{q})$. Поэтому мы не будем останавливаться на обсуждении вида уравнения для матрицы плотности, или основного кинетического уравнения, а сосредоточим внимание на другой аналогии, а именно на аналогии с моментами. Примеры моментов приведены в табл. 13.1 (вторая и

Таблица 13.1. Аналогии между классической и квантовомеханической понятиями

Классическое понятие	Квантовомеханическое понятие
Наблюдаемая q_i	Оператор q_i
Амплитуда поля $E(x)$ в точке с пространственными координатами x	Оператор $E(x)$
Волновая функция электрона $\psi(x)$	Оператор $\psi(x)$
Плотность электронов $\tilde{\rho}(x) = \psi^*(x)\psi(x)$	Оператор $\tilde{\rho}(x) = \psi^+(x)\psi(x)$
Функция распределения $f(q)$	Матрица плотности $\rho = (\rho_{ij})$
Моменты:	
$f_i^{(1)} = \langle q_i \rangle = \int f(\mathbf{q}) q_i d^N q$,	$\langle q_i \rangle = \mathrm{Tr}\{q_i \rho\}$,
$f_{ij}^{(2)} = \langle q_i q_j \rangle = \int f(q) q_i q_j d^N q$,	$Tr\{q_i q_j \rho\}$,
$i = -\int f \ln f d^N q$.	$i = -\mathrm{Tr}\{\rho \ln \rho\}$

третья формулы снизу в левой половине табл. 13.1). Соответствующие выражения в правой части показывают, каким образом аналогичные моменты определены в квантовой теории. Символ Tr означает след. Для матрицы A с элементами a_{ik} след определяется выражением

$$\mathrm{Tr}\{A\} = \sum_j a_{jj}. \tag{13.1}$$

Трудность, возникающая при попытке перевести результаты классической физики на язык квантовой механики, связана с тем, что операторы q_i, вообще говоря, не коммутируют, т.е. в общем случае

$$q_i q_j - q_j q_i \neq 0, \qquad i \neq j. \tag{13.2}$$

Как хорошо известно из квантовой теории, операторы q_i должны в конце концов действовать на волновую функцию. Но в зависимости от последовательности, в которой мы действуем на волновую функцию операторами, окончательные результаты (функции) получаются различными. Именно поэтому столь важно следить за последовательностью, в которой встречаются в квантовой теории операторы. В большинстве рассматриваемых нами приложений мы будем использовать симметризованное произведение q_i и q_j, которое делает несущественной последовательность операторов. Формальное сходство классических и квантовомеханических выражений (табл. 13.1) наводит на мысль о том, что информацию в квантовом случае можно определить так, как это сделано в последней строке табл. 13.1.

После этих предварительных замечаний мы можем обратиться к квантовомеханической формулировке принципа максимума информации.

13.2. Принцип максимума информации

В этом разделе мы воспользуемся схемой соответствия между понятиями, которая приведена в табл. 13.1. Обозначим через S энтропию, или информацию. Отдельные этапы по дальнейшему переводу принципа максимума информации сведены в единую схему в табл. 13.2. Потребуем теперь, чтобы максимума достигало не только выражение для S в классическом случае, но и приведенное в табл. 13.1 соответствующее квантовомеханическое выражение (оба выражения – при соответствующих ограничениях). Ограничения приведены во второй горизонтальной рубрике табл. 13.2, а нормировки – в третьей горизонтальной рубрике той же таблицы. В квантовом случае мы так же, как в классике, при нахождении максимума используем множители Лагранжа. Умножим выражение во второй рубрике на λ_k, выражение в третьей строке – на $\lambda - 1$, вычтем полученные произведения из S и

Таблица 13.2. Принцип максимума информационной энтропии

Классическая система	Квантовомеханическая система
$S = -\int f \ln f d^N q = \max!$	$S = -\mathrm{Tr}\{\rho \ln \rho\} = \max!$
Ограничения	
$f^{(k)} = \int f g^{(k)}(\mathbf{q}) dV$	$f^{(k)} = \mathrm{Tr}\{g^{(k)}(\mathbf{q})\rho\}$
Нормировка	
$\int f d^N q = 1$	$\mathrm{Tr}\{\rho\} = 1$

потребуем, чтобы вариация полученного выражения была равна нулю:

$$\delta [\mathrm{Tr}\{\rho \ln \rho\} - (\lambda - 1)\mathrm{Tr}\{\rho\} - \sum_k \lambda_k \mathrm{Tr}\{g^{(k)}\rho\}] = 0. \qquad (13.3)$$

При варьировании выражения, заключенного в квадратные скобки, необходимо соблюдать осторожность, так как вариация $\delta\rho$, вообще говоря, не коммутирует с ρ:

$$[\rho, \delta\rho] \equiv \rho\,\delta\,\rho - \delta\,\rho\,\rho \neq 0. \qquad (13.4)$$

С другой стороны, мы можем воспользоваться одним важным свойством следа, а именно циклическим свойством следа:

$$\mathrm{Tr}\{AB\} = \mathrm{Tr}\{BA\}. \qquad (13.5)$$

Чтобы проварьировать выражение (13.3), рассмотрим отдельные его члены, начиная с

$$\delta\,\mathrm{Tr}\{\rho\}. \qquad (13.6)$$

Взятие вариации означает, что мы сравниваем значение $\mathrm{Tr}\,\rho$, взятое при значении ρ матрицы плотности, со значением, взятым при другом значении $\rho + \delta\rho$ ($\delta\rho$ — малая величина) матрицы плотности. Таким образом, вариация (13.6) определяется выражением

$$\delta\,\mathrm{Tr}\{\rho\} = \mathrm{Tr}\{\rho + \delta\rho\} - \mathrm{Tr}\{\rho\} = \mathrm{Tr}\{\delta\rho\}, \qquad (13.7)$$

где при получении последнего равенства мы воспользовались линейностью операции взятия следа. Аналогичным образом находим

$$\delta\,\mathrm{Tr}\{g^{(k)}\rho\} = \mathrm{Tr}\{g^{(k)}\delta\rho\}. \qquad (13.8)$$

Вариация первого члена в (13.3) требует известной осторожности

из-за отсутствия коммутативности в общем случае (см. соотношение (13.4)). Пользуясь тем, что вариация следа равна разности следов для $\rho + \delta\rho$ и ρ, получаем

$$\delta\,\mathrm{Tr}\,\{\rho\ln\rho\} = \mathrm{Tr}\,\{(\rho + \delta\rho)\ln(\rho + \delta\rho) - \rho\ln\rho\}. \tag{13.9}$$

Прибавляя и вычитая один и тот же член, преобразуем выражение (13.9) в следующее:

$$(13.9) = \mathrm{Tr}\,\{(\rho + \delta\rho)\ln(\rho + \delta\rho) - \rho\ln(\rho + \delta\rho)\} + \\ + \mathrm{Tr}\,\{\rho\ln(\rho + \delta\rho) - \rho\ln\rho\}. \tag{13.10}$$

Первую разность вычисляем без труда, она оказывается равной

$$\mathrm{Tr}\,\{\delta\rho\ln(\rho + \delta\rho)\} = \mathrm{Tr}\,\{\delta\rho\ln\rho\}, \tag{13.11}$$

где мы удержали только члены, линейные по $\delta\rho$. Чтобы вычислить вторую разность, воспользуемся искусственным приемом, а именно запишем $\ln\rho$ в таком виде, в каком логарифм принято разлагать в ряд Тейлора:

$$\ln\rho = \ln(1 + \rho - 1) = \sum_{\nu=0}^{\infty} c_\nu(\rho - 1)^\nu. \tag{13.12}$$

Совершенно аналогичным образом приходим к разложению

$$\ln(\rho + \delta\rho) = \sum_{\nu=0}^{\infty} c_\nu(\rho - 1 + \delta\rho)^\nu. \tag{13.13}$$

Рассмотрим его отдельные члены. При $\nu = 1$ получаем

$$\rho - 1 + \delta\rho. \tag{13.14}$$

При $\nu = 2$ соответствующий член равен

$$(\rho - 1)^2 + (\rho - 1)\delta\rho + \delta\rho(\rho - 1) + (\delta\rho)^2 \tag{13.15}$$

(с особым вниманием мы следим за тем, чтобы не нарушить последовательность операторов $\rho - 1$ и $\delta\rho$). Общий (ν-й) член разложения имеет следующий вид (мы выписываем только его главные члены):

$$(\rho - 1)^\nu + (\rho - 1)^{\nu-1}\delta\rho + (\rho - 1)^{\nu-2}\delta\rho(\rho - 1) + \\ + (\rho - 1)^{\nu-3}\delta\rho(\rho - 1)^2 + \ldots + \delta\rho(\rho - 1)^{\nu-1} + (\rho - 1)^{\nu-2}(\delta\rho)^2 + \ldots. \tag{13.16}$$

В дальнейшем мы будем удерживать только члены, не зависящие от $\delta\rho$ или линейные относительно $\delta\rho$. В этом приближении умножим выражение (13.16) слева на ρ и возьмем след.

Чтобы продемонстрировать наиболее существенные моменты вычислений, проделаем указанные операции на общем члене выражения (13.16):

$$\mathrm{Tr}\,\{\rho(\rho-1)^{\nu-\mu-1}\,\delta\rho\,(\rho-1)^{\mu}\}. \tag{13.17}$$

Мы можем воспользоваться циклическим свойством следа (см. (13.5)). Тогда

$$(13.17) = \mathrm{Tr}\,\{\delta\rho\,(\rho-1)^{\mu}\,\rho\,(\rho-1)^{\nu-\mu-1}\}. \tag{13.18}$$

Наконец, мы можем воспользоваться тем, что ρ коммутирует с $\rho - 1$, и записать правую часть равенства (13.18) в виде

$$\mathrm{Tr}\,\{\delta\rho\,\rho\,(\rho-1)^{\nu-1}\}. \tag{13.19}$$

Проделанное преобразование позволяет нам прийти к следующему выводу: при взятии следа результат оказывается таким же, как если бы ρ и $\delta\rho$ коммутировали. Это позволяет нам в дальнейшем при вычислении

$$\mathrm{Tr}\,\{\rho\ln(\rho+\delta\rho) - \rho\ln\rho\} \tag{13.20}$$

опускать все алгебраические детали. Мы можем теперь предположить, что ρ и $\delta\rho$ с самого начала коммутируют, или, иначе говоря, обращаться при вычислении выражения (13.20) с операторами ρ и $\delta\rho$ как с *c*-числами (т. е. как с обычными числами, а не с операторами). Используя обычное свойство логарифма, мы преобразуем выражение (13.20) к следующему виду:

$$\mathrm{Tr}\,\{\rho\ln(1+\delta\rho\,\rho^{-1})\}, \tag{13.21}$$

или, если удерживать только главный член,

$$(13.20) = \mathrm{Tr}\,\{\delta\rho\}. \tag{13.22}$$

Используя соотношения (13.7), (13.8), (13.20) и (13.22), вычисляем вариацию (13.3) и получаем

$$-\,\mathrm{Tr}\,\{\delta\rho\ln\rho\} - \lambda\,\mathrm{Tr}\,\{\delta\rho\} - \sum_k \lambda_k\,\mathrm{Tr}\,\{\delta\rho\, g^{(k)}\} = 0. \tag{13.23}$$

Чтобы реально проварьировать ρ, т. е. получить $\delta\rho$, можно воспользоваться любым представлением ρ относительно множества собственных состояний с дискретным или непрерывным спектром. Такое представление позволяет записать ρ, $\delta\rho$ и $g^{(k)}$ в виде матриц.

Выразим теперь различные члены соотношения (13.23) с учетом определения следа:

$$\sum_j (\delta\rho\ln\rho)_{jj} = \sum_{j,\,l} (\delta\rho)_{jl}(\ln\rho)_{lj}, \tag{13.24}$$

$$\sum_j (\delta\rho)_{jj}, \tag{13.25}$$

$$\sum_{j,\,l} (\delta\rho)_{jl}\,(g^{(k)})_{lj}. \tag{13.26}$$

Подставляя соответствующие выражения в соотношение (13.23), получаем

$$\sum_j (\delta\rho)_{jl}\,[(\ln\rho)_{lj} + \lambda\,\delta_{lj} + \sum_k \lambda_k (g^{(k)})_{lj}] = 0, \tag{13.27}$$

где мы воспользовались символом Кронекера δ_{lj}, равным 1 при $l = j$ и 0 при $l \neq j$. Так как мы используем множители Лагранжа, вариации $(\delta\rho)_{jl}$ можно считать независимыми друг от друга. Это означает, что соотношение (13.27) может выполняться только в том случае, если выражение, заключенное в квадратные скобки, равно нулю при любых двух значениях индексов l и j. Но каждое такое выражение можно рассматривать как элемент некоторой матрицы с индексами l и j. Таким образом, обращение в нуль выражений, стоящих в квадратных скобках в левой части соотношения (13.27), эквивалентно матричному уравнению

$$\ln\rho + \lambda I + \sum_k \lambda_k g^{(k)} = 0, \tag{13.28}$$

где I – единичная матрица. Матричное или операторное уравнение (13.28) имеет решение

$$\rho = \exp\left(-\lambda - \sum_k \lambda_k g^{(k)}\right). \tag{13.29}$$

Множители Лагранжа λ и λ_k (по крайней мере в принципе) могут быть определены из условия нормировки

$$\mathrm{Tr}\,\{\rho\} = 1 \tag{13.30}$$

и ограничений

$$\mathrm{Tr}\,\{g^{(k)}\rho\} = f^{(k)}. \tag{13.31}$$

Подставляя решение (13.29) в условие (13.30), получаем

$$\mathrm{Tr}\,\{\rho\} = e^{-\lambda}\,\mathrm{Tr}\,\{\exp\left(-\sum_k \lambda_k g^{(k)}\right)\} = 1. \tag{13.32}$$

Из формулы (13.32) сразу же следует соотношение

$$e^{\lambda} = Z = \mathrm{Tr}\,\{\exp\left(-\sum_k \lambda_k g^{(k)}\right)\}, \tag{13.33}$$

которое, кроме того, содержит определение квантовомеханической функции распределения Z. Ряд соотношений, выведенных нами в разд. 3.3, допускают обобщение на квантовомеханический случай. Но, как увидит читатель, это обобщение требует известной осторожности и некоторых искусственных приемов. Мы не будем приводить здесь полностью вывод квантовомеханических соотношений, а лишь

укажем основные этапы вывода и приведем окончательный результат. Чтобы вывести квантовомеханические аналоги соотношений (3.44) и (3.45), нам необходимо знать, как дифференцировать экспоненциальную функцию, содержащую операторы, которые, вообще говоря, не коммутируют. Так как остальная часть этого раздела носит сугубо технический характер, читатель, которого не интересуют подробности, может, опустив все промежуточные результаты, ознакомиться с окончательной формулой (13.43) и затем перейти к следующему разделу.

Сформулируем цель наших выкладок. Мы хотим вычислить

$$\frac{\partial}{\partial \lambda_{k_0}} \mathrm{Tr}\{\rho\}, \tag{13.34}$$

т. е. хотим научиться вычислять производную от матрицы плотности, стоящей под знаком взятия следа. Для этого рассмотрим экспоненту, стоящую в левой части равенства .

$$\exp(A + \lambda' B) = T\exp\left[\int_0^1 dt (A_t + \lambda' B_t)\right]. \tag{13.35}$$

Эта экспонента содержит тот самый параметр λ', по которому нам требуется продифференцировать экспоненциальную функцию. Но поскольку операторы A и B, вообще говоря, не коммутируют, операция дифференцирования не вполне обычна. Чтобы выполнить ее, мы должны прибегнуть к особому приему – так называемому хронологическому упорядочению операторов. Присвоим операторам A и B индекс t и заменим левую часть равенства (13.35) выражением, стоящим в правой части того же равенства. Здесь T означает хронологическое упорядочение: при вычислении экспоненты операторы с меньшим значением индекса t должны предшествовать операторам с бóльшим значением индекса t. При соблюдении этого соглашения мы можем обращаться с операторами A_t и B_t так, как если бы эти операторы коммутировали друг с другом. Таким образом, соотношение (13.35) может быть записано в виде

$$T\exp\left(\int_0^1 dt\, A_t\right)\exp\left(\int_0^t \lambda' dt\, B_t\right). \tag{13.36}$$

Дифференцируя равенство (13.35) или, что эквивалентно, выражение (13.36) по λ', т. е. вычисляя производную

$$\frac{d}{d\lambda'} T\exp\left[\int_0^1 dt (A_t + \lambda' B_t)\right], \tag{13.37}$$

получаем

$$T\int_0^1 B_{t'}\, dt' \exp\left[\int_0^1 dt (A + \lambda' B)_t\right]. \tag{13.38}$$

Принятое нами соглашение о хронологическом упорядочении означает, что некоторые операторы, стоящие под знаком экспоненты, а именно те, которые соответствуют значениям индекса $t < t'$, должны предшествовать оператору $B_{t'}$. С другой стороны, операторы, стоящие под знаком экспоненты и соответствующие значениям индекса $t' < t$, должны следовать за оператором $B_{t'}$. Таким образом, мы производим хронологическое упорядочение, разбивая соответствующим образом экспоненту, и получаем

$$\int_0^1 dt' \exp\left[\int_{t'}^1 dt\left(A + \lambda' B\right)_t\right] B_{t'} \exp\left[\int_0^{t'} dt\left(A + \lambda' B\right)_t\right]. \tag{13.39}$$

Выполнив хронологическое упорядочение под знаком каждой из экспонентов в отдельности, мы можем заменить операторы A_t и B_t операторами без индекса и произвести интегрирование. В результате мы получаем

$$\int_0^1 dt' \exp\left[(1 - t')(A + \lambda' B)\right] B \exp\left[t'(A + \lambda' B)\right]. \tag{13.40}$$

Если взять след оператора (13.40) и воспользоваться циклическим свойством (13.5), то первую экспоненту в выражении (13.40) можно перенести вправо. Две потенциальные функции можно слить в одну, которая не будет зависеть от t'. После этого интегрирование по t' проводится без труда, и мы получаем единицу. Замечая, что операции дифференцирования и взятия следа можно производить в любом порядке, приходим к требуемому соотношению

$$\frac{d}{d\lambda'} \mathrm{Tr}\left\{\exp(A + \lambda' B)\right\} = \mathrm{Tr}\left\{B \exp(A + \lambda' B)\right\}. \tag{13.41}$$

Напомним нашу первоначальную цель: мы хотели продифференцировать след матрицы плотности ρ по λ_k. Стоит для этого теперь отождествить B с $g^{(k)}$, а остальную часть экспоненты от ρ – с A, как мы сразу получаем

$$\frac{\partial}{\partial \lambda_k} \mathrm{Tr}\{\rho\} = \mathrm{Tr}(g^{(k)} \rho). \tag{13.42}$$

Наконец, используя явное выражение для ρ и определение Z, приходим к окончательному результату:

$$f_k = \mathrm{Tr}\{g^{(k)} \rho\} = -\frac{\partial}{\partial \lambda_k} \ln Z, \tag{13.43}$$

который полностью аналогичен формуле (3.45). Если нам понадобится вычислить вторые производные от Z по λ_k, то некоторые из примененных нами трюков окажутся полезными, но окончательный

результат будет выглядеть несколько сложнее, чем соответствующие результаты из разд. 3.3.

В заключение раздела заметим, что изложенные в гл. 4 термодинамические результаты также могут быть перенесены на квантовомеханический случай. Например, если в качестве ограничения выбрана энергия системы, то

$$\rho = Z^{-1}\exp\left(-\frac{H}{kT}\right), \tag{13.44}$$

где H – гамильтониан, k – постоянная Больцмана и T – абсолютная температура. Функция распределения Z определяется соотношением

$$Z = \mathrm{Tr}\left\{\exp\left(-\frac{H}{kT}\right)\right\}. \tag{13.45}$$

13.3. Параметры порядка, подчиненные моды и структуры

Здесь мы хотим выяснить, каким образом результаты, полученные в гл. 6, могут быть перенесены на квантовомеханические системы. Для этого мы опять особым образом выберем ограничения

$$f_k = \mathrm{Tr}\{g^{(k)}\rho\}, \tag{13.46}$$

а именно заменим f_k симметризованными моментами

$$\langle q_j\rangle, \tag{13.47}$$

$$\langle q_j q_k + q_k q_l\rangle \tag{13.48}$$

и т. д. В дальнейшем мы будем включать в число ограничений моменты до четвертого порядка включительно. При этом матрица плотности, максимизирующая информацию, определяется выражением

$$\rho = \exp\left(-\lambda - \sum_j \lambda_j q_j - \sum_{j,j'} \lambda_{jj'} q_j q_{j'} - \ldots\right), \tag{13.49}$$

где показатель экспоненты содержит члены до четвертого порядка по q_j включительно. В дальнейшем условимся для краткости обозначать показатель через V, т. е. полагать

$$\rho = e^V. \tag{13.50}$$

Мы будем как можно более точно следовать процедуре, описанной в гл. 6. Прежде всего требуется исключить из показателя выражений (13.49) и (13.50) линейные члены. Для этого мы предполагаем, что

$$q_j = c_j + \bar{q}_j, \tag{13.51}$$

где c_j – это c-числа. Затем мы подставляем разложение (13.51) в

формулу (13.49) и получаем, в частности, члены, линейные по q_j, коэффициенты которых должны быть равны нулю:

$$[\lambda_j + \sum_{j'} (\lambda_{jj'} c_{j'} + \lambda_{j'j} c_{j'}) + \ldots] = 0. \tag{13.52}$$

Как нетрудно убедиться, к соотношению (13.52) мы приходим также, потребовав, чтобы выполнялось условие

$$\frac{\partial V}{\partial q_i} = 0 \qquad \text{при} \qquad q_j^0, \tag{13.53}$$

где q_j мы рассматриваем как c-числа.

Поскольку мы отождествляем

$$q_j^0 \leftrightarrow \bar{q}_j = 0, \tag{13.54}$$

коэффициенты c_j определяются равенствами

$$c_j = q_j^0. \tag{13.55}$$

В результате мы преобразуем правую часть формулы (13.49) к следующему виду:

$$\exp[-\bar{V}(\bar{\mathbf{q}})], \tag{13.56}$$

где вместо V теперь стоит показатель

$$\bar{V}(\bar{\mathbf{q}}) = \bar{\lambda} + \sum_{j,j'} \bar{\lambda}_{jj'} \bar{q}_j \bar{q}_{j'} + \ldots . \tag{13.57}$$

Поскольку мы выбрали симметризованные моменты, можно быть уверенными в том, что

$$\bar{\lambda}_{jj'} = \bar{\lambda}_{j'j}. \tag{13.58}$$

Предположим теперь, что

$$\bar{q}_j = \sum_k \xi_k v_{kj}, \tag{13.59}$$

где v_{kj} есть c-числа, в то время как ξ_k – операторы. Подставляя соотношение (13.59) в показатель (13.57), получаем для квадратичных членов выражение

$$\sum_{k,k'} \xi_k \xi_{k'} \left\{ \sum_{j,j'} \bar{\lambda}_{jj'} v_{kj} v_{k'j'} \right\}. \tag{13.60}$$

Выберем коэффициенты v_{kj} так, чтобы выполнялось условие

$$\sum_{j,j'} \bar{\lambda}_{jj'} v_{kj} v_{k'j'} = \tilde{\lambda}_k \delta_{kk'}. \tag{13.61}$$

Поскольку $\bar{\lambda}_{jj'}$ – симметричная матрица, коэффициенты v_{kj} можно выбрать так, чтобы их матрица и значения $\hat{\lambda}_k$ были вещественны. В результате таких преобразований формула (13.49) переходит в соотношение

$$\rho(\xi_u, \xi_s) = \exp\left[-\bar{\lambda} - \sum_u \hat{\lambda}_u \xi_u^2 - \sum_s \hat{\lambda}_s \xi_s^2 + V_u(\xi_u) + V_s(\xi_u, \xi_s)\right], \quad (13.62)$$

где мы снова ввели особые обозначения для положительных ($\hat{\lambda}_u$) и отрицательных ($\hat{\lambda}_s$) значений $\hat{\lambda}$ для того, чтобы можно было легко отличать параметры порядка от подчиненных мод.

Введем теперь матрицу плотности как функцию одних лишь параметров порядка ξ_u с помощью соотношения

$$\rho_u(\xi_u) = \mathrm{Tr}_s\{\rho(\xi_u, \xi_s)\}. \quad (13.63)$$

Так как полный след нормирован, справедливо равенство

$$\mathrm{Tr}\{\rho_u\} = 1. \quad (13.64)$$

По аналогии с процедурой, изложенной в гл. 6, мы хотим теперь разложить матрицу плотности ρ, которая зависит от операторов параметров порядка и подчиненных мод, в произведение матрицы плотности (13.63) и своего рода условной матрицы плотности по формуле

$$\rho(\xi_u, \xi_s) = \rho_u(\xi_u)\,\rho_s(\xi_s \mid \xi_u). \quad (13.65)$$

Взяв след от обеих частей последнего равенства по переменным, относящимся к операторам ξ_s, мы по определению получим в левой части величину (13.63). Правая часть равна

$$\rho_u(\xi_u)\,\mathrm{Tr}_s\{\rho_s(\xi_s \mid \xi_u)\}. \quad (13.66)$$

Сравнивая правую и левую части, получаем соотношение

$$\mathrm{Tr}_s\{\rho_s(\xi_s \mid \xi_u)\} = 1. \quad (13.67)$$

Мы хотим теперь преобразовать матрицу плотности (13.62) так, чтобы ее можно было разложить в произведение (13.65). Это отнюдь не тривиальная задача, так как операторы ξ не коммутируют между собой. Чтобы решить ее, запишем матрицу плотности (13.62) в виде

$$\rho(\xi_u, \xi_s) = \exp[\hat{V}_u(\xi_u) + \hat{V}_s(\xi_u, \xi_s)], \quad (13.68)$$

где обозначения очевидны, а вопрос о том, входит ли $\bar{\lambda}$ в V_u или в V_s, оставлен открытым.

Прибавляя к показателю экспоненты операторную функцию h и вычитая ее, получаем

$$(13.68) = \exp[\hat{V}_u + h(\xi_u) - h(\xi_u) + \hat{V}_s(\xi_u, \xi_s)]. \quad (13.69)$$

Наша задача состоит в том, чтобы записать матрицу плотности (13.69) в виде произведения (13.65). Для этого нам понадобится лемма. Рассмотрим выражение

$$T\exp\left[\int_0^1 dt\,(A_t + B_t)\right], \tag{13.70}$$

которое получается из соотношения (13.35) при $\bar{\lambda} = 1$. Здесь T по-прежнему означает оператор хронологического упорядочения, а A_t и B_t — операторы, подлежащие хронологическому упорядочению. Воспользуемся теоремой Фейнмана о разделении операторов. По этой теореме оператор (13.70) может быть представлен в виде

$$e^A\, Te^{\tilde{B}}, \tag{13.71}$$

где

$$\tilde{B} = \int_0^1 dt\,(e^{-At} B e^{At})_t. \tag{13.72}$$

Пусть теперь

$$\hat{V}_u + h = A; \qquad -h + \hat{V}_s = B. \tag{13.73}$$

Тогда

$$\rho_u(\xi_u) = \exp[\hat{V}_u + h], \tag{13.74}$$

в то время как условная матрица плотности определяется соотношением

$$\rho_s(\xi_s \mid \xi_u) = Te^{\tilde{B}}. \tag{13.75}$$

Неизвестную пока операторную функцию h следует определить так, чтобы

$$\mathrm{Tr}_s\{\rho_s(\xi_s \mid \xi_u)\} = 1. \tag{13.76}$$

Попробуем теперь подвести некоторые итоги. Начав с операторных моментов до четвертого порядка включительно, мы с помощью принципа максимума информационной энтропии построили матрицу плотности. Затем, диагонализуя ее билинейную часть, мы выделили особо операторы, соответствующие параметрам порядка, и особо — операторы, соответствующие подчиненным модам. Наконец, нам удалось записать совместную матрицу плотности для ξ_u, ξ_s в виде произведения, где множитель ρ_u — матрица плотности, относящаяся только к параметрам порядка, в то время как множитель ρ_s можно интерпретировать как условную матрицу плотности. Совершенно аналогично тому, как это делается в классическом случае, мы можем интерпретировать v_{kj} как структуру в пространственных точках j, принадлежащих моде k, если q_j, например, интенсивность, соответствующая пространственной точке j. Следует особо отметить, что

v – классическая величина. Это означает, что возникающие структуры описываются классическими величинами, и, следовательно, их интерпретация не требует знания никакой квантовомеханической волновой функции.

В заключение этой главы мы хотим показать, что информация допускает разбиение на две части, одна из которых содержит только параметры порядка, а другая – условную матрицу плотности, т. е. подчиненные моды.

13.4. Информация, содержащаяся в параметрах порядка и в подчиненных модах

Начнем с выражения для информации всей системы

$$i = \mathrm{Tr}\{\rho \ln \rho\} \tag{13.77}$$

и подставим в него условную матрицу плотности (13.65), в которой два сомножителя определены соответственно соотношениями (13.74) и (13.75). Учитывая, кроме того, формулу (13.69), получаем

$$(13.77) = \mathrm{Tr}\{\rho_u \rho_s [(\hat{V}_u + h) + (-h + \hat{V}_s)]\}. \tag{13.78}$$

Правая часть последнего выражения представима в виде суммы двух членов:

$$\mathrm{Tr}_u\{\rho_u(\mathrm{Tr}_s\{\rho_s\})(\hat{V}_u + h)\} \tag{13.79}$$

и

$$\mathrm{Tr}_u\{\rho_u \mathrm{Tr}_s\{\rho_s(-h + \hat{V}_s)\}\}. \tag{13.80}$$

Соотношение (13.67) позволяет упростить выражение (13.79) и свести его к следующему:

$$\mathrm{Tr}_u\{\rho_u(\hat{V}_u + h)\}, \tag{13.81}$$

в то время как след по s в выражении (13.80) может быть интерпретирован как информация, соответствующая подчиненной моде s. Поскольку эта информация зависит от операторов ξ_u, мы имеем дело с операторным выражением для информации и можем записать соотношение

$$i_{s,\,\text{опер}}(\xi_u) = \mathrm{Tr}_s\{\rho_s(-h + \hat{V}_s)\}. \tag{13.82}$$

Обозначая выражение (13.81) для краткости просто через i_u, получаем для i (см. формулу (13.77)) окончательное выражение

$$i = i_u + \mathrm{Tr}_u\{\rho_u i_{s,\,\text{опер}}(\xi_u)\}. \tag{13.83}$$

Таким образом, совершенно аналогично тому, как это происходит в классическом случае, мы можем разделить информацию всей системы на две части – информацию, относящуюся к параметрам порядка, и

информацию, относящуюся к подчиненным модам, которая затем подлежит усреднению по функции распределения или матрице плотности параметров порядка.

И наконец, мы хотели бы еще раз заявить, что очень многие результаты, полученные ранее в рамках классической теории, допускают обобщение на случай квантовомеханических операторов. В частности, предлагаемая нами процедура позволяет идентифицировать структуры, возникающие в квантовых системах, с помощью матриц v_{kj}. Это означает возможность распознавания классических структур в квантовых системах.

Глава 14
Заключительные замечания и некоторые перспективы

В этой главе, которая завершает нашу книгу, я хочу кратко обсудить полученные результаты и обратить внимание читателя на те области, в которых еще остаются нерешенные вопросы.

Глава 1 посвящена природе сложных систем. Насколько можно судить, основная отличительная особенность сложных систем – их поистине неисчерпаемое разнообразие, в котором нам так и не удается разобраться, несмотря на многочисленные попытки понять до конца природу сложных систем. Это вынуждает нас ограничиваться рассмотрением тех отдельных аспектов сложных систем, которые по тем или иным соображениям представляются нам важными или интересными. В этой книге мы попытались справиться с разнообразием сложных систем, сосредоточив внимание на одной, весьма общей их особенности, а именно на способности сложных систем при определенных условиях качественно изменять свое макроскопическое поведение. Точнее говоря, мы рассмотрели так называемые неравновесные фазовые переходы. Избранный нами подход позволяет анализировать функционирование не только таких физических систем, как лазеры или жидкости, но гораздо более сложных, например биологических, систем. Нет никаких сомнений в том, что круг приложений нашего подхода может и будет существенно расширен. Как показывает пример лазера, предлагаемый нами подход позволяет также рассматривать колебательные явления, хотя вопрос о его применимости к хаотическим системам остается пока открытым. Однако нет никаких сомнений в том, что здесь открывается широкий простор для исследований. В основе предлагаемого нами метода лежит понятие информации. Как было показано, термин «информация» может употребляться в различных смыслах, поэтому мы начали с обсуждения различных определений этого термина. В частности, мы показали, что понятие *информации* в смысле Шеннона в некоторых случаях оказывается более подходящим, чем понятие *энтропии*. Дело в том, что по крайней мере понятие энтропии имеет в физике вполне определенное значение и применимо к системам, находящимся в тепловом равновесии. Если мы рассматриваем физические системы, далекие от теплового равновесия, или биологические системы, то результаты необходимо интерпретировать заново, и для того чтобы максимизировать информацию или (статистическую) энтропию, приходится вводить новые ограничения.

В частности, как было показано, возникает в некотором смысле новый тип информации, связанный с коллективными переменными

или параметрами порядка. Это навело нас на мысль назвать ту часть информации, которая относится к параметрам порядка и отражает коллективные свойства системы, *синергетической информацией.* Что же касается самих параметров порядка, то они обретают новый смысл, превращаясь в носителей информации – «информаторов».

Важно выяснить, какого рода информацию в обычном смысле позволяет обнаружить наш общий подход, основанный на использовании моментов переменных. Как было показано, предлагаемая нами процедура позволяет определять параметры порядка и их стационарные функции распределения. Достаточна ли такого рода информация или недостаточна, зависит от того, какую цель мы ставим перед собой. Рассмотрим в качестве примера лазер. Стационарной функции распределения вполне достаточно для того, чтобы охарактеризовать стационарную интенсивность и флуктуации интенсивности лазерного излучения. Но такая функция распределения ничего не говорит нам об очень важном свойстве лазерного излучения, а именно о его временной когерентности. Чтобы получить информацию об этом свойстве, необходимо гораздо более детальное знание динамики системы. В рамках нашего подхода такого рода динамика может быть восстановлена по зависящим от времени моментам и вероятностям переходов, вычисляемым по известным моментам с помощью принципа максимума информационной энтропии. Обобщая сказанное, можно утверждать, что если известны средние значения каких-то макроскопических данных, то предлагаемая процедура позволяет строить предположения относительно микроскопических событий и процессов, которые кроятся за этими макроскопическими данными.

Трудно удержаться от искушения и не высказать утверждение о том, что в системах, далеких от теплового равновесия, или даже в нефизических системах информация (Шеннона) играет такую же роль, как энтропия в системах, находящихся в тепловом равновесии или близких к нему, а именно является *причиной протекания процессов.* Я склонен придерживаться такой интерпретации по следующим причинам. Прежде всего мы обнаружили, что результаты нашего анализа зависят от надлежаще выбранных ограничений на систему. Нам удалось определить адекватные ограничения для одного класса явлений, которые можно назвать неравновесными фазовыми переходами. Мы можем также строить предположения относительно функции распределения. В этом отношении можно сказать, что мы можем определить потенциал, переводящий систему в определенные стационарные состояния. Но этот потенциал не дает нам однозначного алгоритма, например, для вывода уравнения Фоккера–Планка, определяющего динамику. На это обстоятельство неоднократно указывал Р. Ландауэр. Таким образом, создается впечатление, что максимизация информации (или энтропии) не является фундаментальным законом, который однозначно определяет эволюцию систем. С другой стороны, мы знаем замечательные работы Грэхэма,

Теля и других исследователей, сумевших построить потенциалы для стохастических процессов в тех случаях, когда происходят бифуркации. Мы не предпринимали попыток сравнить наш подход с их подходом. Такое сравнение остается интересной задачей для будущих исследований.

Затронутые нами идеи имеют первостепенное значение для вопросов, относящихся к биологической эволюции, развитию, или, иначе говоря, к филогенезу и онтогенезу животных. Основной вопрос, разумеется, состоит в том, подчиняются ли эволюция и развитие экстремальным принципам, в частности экстремальным принципам, связанным с одной функцией, такой, как энтропия или информация. Наши результаты свидетельствуют о том, что такая функция может существовать, но достается достаточно дорогой ценой: располагая такой функцией, мы не можем ничего сказать о динамике или знаем о ней очень мало. Но и самый вопрос о существовании такой функции остается пока открытым. Вместе с тем то, что выбор моментов дает адекватные ограничения для информации в неравновесных фазовых переходах, показывает силу нашего подхода. Поведение моментов отражает тенденцию системы к переходу в когерентное состояние. Таким образом, в этой книге мы по существу занимались исследованием того, как в сложных системах с помощью самоорганизации возникают когерентность и макроскопический порядок. Но вопрос о том, можно ли биологию рассматривать исключительно с этой точки зрения, остается открытым. Как уже было сказано, сложные системы пока все еще представляются нам неисчерпаемо разнообразными.

Резюмируя, можно утверждать, что, хотя глубокий вопрос, имеющий первостепенное значение, остается отрытым, изложенный в этой книге подход, несомненно, может служить эффективным средством, позволяющим с единой точки зрения исследовать когерентные свойства сложных систем вблизи неравновесных фазовых переходов.

Литература

Глава 1

Разд. 1.5.1 и 1.5.2

Ландау Л. Д., Лифшиц Е. М. Статистическая физика, 3-е изд., ч. 1.– М.: Наука, 1976 (Теоретическая физика, т. 5).
Becker R., Theory of Heat, Springer, Berlin, Heidelberg, 1967.
Münster A., Statistical Thermodynamics, Vol. 1, Springer, Berlin, Heidelberg, 1969.
Callen H. B., Thermodynamics, Wiley, New York, 1960.
Landsberg P. T., Thermodynamics, Wiley, New York, 1961.
Kubo R., Thermodynamics, North Holland, Amsterdam, 1968. [Имеется перевод: *Кубо Р.* Термодинамика. Современный курс с задачами и решениями.– М.: Мир, 1970.]

Разд. 1.5.3

Haken H., Synergetics, An Introduction, Spriger Ser. Synergetics, Vol. 1, 3rd ed., Springer, Berlin, Heidelberg, 1983. [Имеется перевод 1-го изд.: *Хакен Г.* Синергетика.– М.: Мир, 1980.]
Haken H., Advanced Synergetics, Springer Ser. Synergetics, Vol. 20, 2nd corr. printing, Springer, Berlin, Heidelberg, 1987. [Имеется перевод: *Хакен Г.* Синергетика: иерархии неустойчивостей в самоорганизующихся системах и устройствах.– М.: Мир, 1985.]

Разд. 1.6.1

Shannon C. E., A Mathematical Theory of Communication, Bell System Techn. J., **27**, 370–423, 623–656 (1948). [См. также: *Шеннон К.* Работы по теории информации и кибернетике. Сб. перев. с англ.– М.: ИЛ, 1963.]
Shannon C. E., Bell System Techn. J., **30**, 50 (1951).
Shannon C. E., Weaver W., The Mathematical Theory of Communication, Univ. of Illin. Press, Urbana, 1949.
Botzmann L., Vorlesungen über Gastheorie, 2 vols., Leipzig, 1896, 1898.

Разд. 1.6.2

Haken H., In: Thermodynamics and Regulation of Biological Processes, ed. by I. Lamprecht, A. I. Zotin, Walter de Gruyter & Co, Berlin, New York, 1984.

Разд. 1.6.3

Haken H., статья, представленная на международн. конф. по информатике, Токио, 1986.
Haken H., In: Ordnung aus dem Chaos. Prinzipien der Selbstorganisation und Evolution des Lebens, Bernd-Olaf Küppers (Hrsg.), Piper, München, Zürich, 1987.

Глава 2

См. литературу к разд. 1.5.3.

Глава 3

Разд. 3.2

Kulback S., Ann. Math. Statist., **22**, 79 (1951).
Kulback S., Information Theory and Statistics, Wiley, New York, 1951.

Разд. 3.3

Jaynes E. T., Phys. Rev., **106**, 4, 620 (1957); **108**, 171 (1957).
Jaynes E. T., In: Delaware Seminar on the Foundations of Physics, Springer, Berlin, Heidelberg, 1967.
Jaynes E. T., Am. J. Phys., **33**, 391 (1965).
Jaynes E. T., In: The Maximum Entropy Formalism, ed. by R. D. Levine, M. Tribus, MIT Press, Cambridge, Mass., 1978.

Глава 5

Ранние приложения принципа Джейнса к сиситемам с несколькими температурами были даны Р. Ингарденом, его сотрудниками и другими авторами в многочисленных статьях. См., например:

Ingarden R. S., Bull. Acad. Polon. Sci., Ser. Sci. Math. Astron. Phys., **11**, 541 (1963).
Ingarden R. S., Fortschr. Phys., **12**, 567 (1964b); **13**, 755 (1965a).
Ingarden R. S., Acta Phys. Polon., **27**, 179 (1965b).
Ingarden R. S., Ann. Inst. Henri Poincare, **8**, 1 (1968a).
Ingarden R. S., Bull. Acad. Polon. Sci., Ser. Sci. Math. Astron. Phys., **16**, 57 (1968b).
Ingarden R. S., Acta Physica Polonica, Vol. XXXVI, Fasc. **5**(11), 855 (1969).

В данной главе и в гл. 6 я по существу следую работам:

Haken H., Z. Phys. B-Condensed Matter, **61**, 335 (1985).
Haken H., Z. Phys. B-Condensed Matter, **63**, 487 (1986).

Разд. 5.2

Вывод функции распределения для лазера см. в работах:

Risken H., Z. Phys., **186**, 85 (1965).
Haken H., Laser Theory, Corr. printing, Springer, Berlin, Heidelberg, 1984; первая публикация в: Handbuch der Physik, Vol. 25/2c, ed. by S. Flügge, 1970.

Разд. 5.3

Haken H., Laser Theory, Corr. printing, Springer, Berlin, Heidelberg, 1984.

Глава 6

См. литературу к гл. 5.

Глава 7

Здесь вплоть до разд. 7.8 я следую работам:

Haken H., Z. Physik B–Condensed Matter, **61**, 329 (1985).
Haken H., Z. Physik B–Condensed Matter, **62**, 255 (1986).

Разд. 7.9

Tuszynsky J. A., Z. Phys. B–Condensed Matter, **65**, 375 (1987).
Witschel W., Chem. Phys., **50**, 265 (1980).
Witschel W., *Bohmann J.*, J. Phys., **A13**, 2735 (1980).

Разд. 7.10

Здесь я следую по существу работам:

Klimontovich Yu. L, Physica, **A142**, 390 (1987).
Klimontovich Yu. L, Z. Phys. B–Condensed Matter, **65**, 125 (1987).

Глава 8

Разд. 8.2

Haken H., *Weberruss J.*, неопубликованные результаты.

Глава 9

Haken H., Z. Phys. B–Condensed Matter, **63**, 505 (1986).

Глава 10

Разд. 10.1

Haken H., In: Frontiers in Quantum Optics, ed. by E. R. Pike, S. Sarkar, Adam Hilger, Bristol, 1986.
Haken H., Laser Theory, Corr. printing, Springer, Berlin, Heidelberg, 1984; первоначально издано в: Handbuch der Physik, Vol. 25/2c, ed. by S. Flugge, 1970.

Разд. 10.2

Haken H., In: Frontiers in Quantum Optics, ed. by E. R. Pike, S. Sarkar, Adam Hilger, Bristol, 1986.

Глава 11

Эксперименты см.:

Kelso J. A. S., Bull. Psychon. Soc., **18**, 63 (1981).
Kelso J. A. S., Am. J. Phys.: Reg. Integr. Comp., **15**, R1000–1004 (1984).

Теория (в терминах детерминистических или стохастических уравнений) была развита в работах:

Haken H., Kelso J. A. S., Bunz H. H., Biol. Cybern., **51**, 347 (1985).
Schoner G., Haken H., Kelso J. A. S., Biol. Cybern., **53**, 247 (1986).

Глава 12

Литература по распознаванию образцов компьютером растет взрывообразно; см., например:

Haken H. (ed.), Pattern Formation by Dynamic Systems and Pattern Recognition, Springer Ser. Syn., Vol. 5, Springer, Berlin, Heidelberg, 1979.
Denker J. S. (ed.), Neural Networks for Computing, AIP Conf. Proc., 151, New York, 1986.
Candill M., Butler S., IEEE First Int. Conf. on Neural Networks, Vols. I–IV, SOS Printing, San Diego, 1987.

Разд. 12.1

Devijer P. A., Kittler J., Pattern Recognition, A Statistical Approach, Prentice Hall, Englewood Cliffs, N. J., 1982.

Разд. 12.2

Haken H., In: Computational Systems, Natural and Artificial, ed. by H. Haken, Springer Ser. Syn., Vol. 38, Springer, Berlin, Heidelberg, 1987.

Разд. 12.3

Haken H., не опубликовано.

Разд. 12.4

Haken H., не опубликовано.

Специальный случай спиновых стекол см.:

Ackley D. H., Hinton G. E., Sejnovski, Cognitive Science, **9**, 147 (1985).

Разд. 12.5

Haken H., не опубликовано.

Глава 14

Graham R., Tel T., Phys. Rev., **35A**, 1328 (1987).
Graham R., Tel T., Phys. Rev. Lett., **52**, 9 (1984).
Graham R., Tel T., J. Stat. Phys., **35**, 729 (1984); **37**, 709 (1984).
Graham R., Tel T., Phys. Rev., **A31**, 1109 (1985).
Graham R., Roekaerts D., Tel T., Phys. Rev., **A31**, 3364 (1985).
Freidlin M. I., Wentzell A. D., Random Perturbations of Dynamical Systems, Springer, New York, 1984.
Landauer R., Relative Stability in the Dissipative State, Preprint, 1978.

Оглавление

Предисловие проф. Ю. Л. Климонтовича 5

Предисловие к русскому изданию 11

Предисловие 12

Глава 1. Сложные системы – вызов искусству исследователя 14

1.1. Что такое сложные системы? 14
1.2. Как подходить к исследованию сложных систем? 21
1.3. Модельные системы 23
1.4. Самоорганизация 28
1.5. В поисках универсальности 29
1.5.1. Термодинамика 29
1.5.2. Статистическая физика 31
1.5.3. Синергетика 31
1.6. Информация 33
1.6.1. Шенноновская информация: изгнание смыслового содержания 34
1.6.2. Как информация действует на систему 35
1.6.3. Саморождение смысла 45
1.6.4. Сколько информации необходимо для поддержания упорядоченного состояния? 54
1.7. Второе начало синергетики 58

Глава 2. От микроскопического мира к макроскопическому 63

2.1. Уровни описания 63
2.2. Уравнения Ланжевена 65
2.3. Уравнение Фоккера–Планка 67
2.4. Точное стационарное решение уравнения Фоккера–Планка для систем, находящихся в детальном равновесии 68
2.4.1. Детальное равновесие 69
2.4.2. Требуемая структура уравнения Фоккера–Планка и его стационарное решение 70
2.5. Интегралы по траекториям 71
2.6. Уменьшение сложности, параметры порядка и принцип подчинения 73
2.6.1. Анализ устойчивости по линейному приближению 74
2.6.2. Преобразование эволюционных уравнений 75
2.6.3. Принцип подчинения 76
2.7. Неравновесные фазовые переходы 77
2.8. Образование структур 79

Глава 3. ... и снова назад: принцип максимума информации (ПМИ) 81

3.1. Некоторые основные идеи 81
3.2. Приращение информации 86
3.3. Информационная энтропия и ограничения, налагаемые на систему 87
3.4. Непрерывные переменные 93

Глава 4. Пример из области физики: термодинамика 94

Глава 5. Приложение принципа максимума информации к самоорганизующимся системам 99

5.1. Введение 99
5.2. Приложение к самоорганизующимся системам: одномодовый лазер 99
5.3. Многомодовый лазер без фазовых соотношений 101
5.4. Процессы, периодические по параметрам порядка 102

Глава 6. Принцип максимума информации для неравновесных фазовых переходов: определение параметров порядка, подчиненные моды и возникающие структуры 105

6.1. Введение 105
6.2. Общий подход 105
6.3. Определение параметров порядка, подчиненных мод и возникающих структур 107
6.4. Приближения 109
6.5. Пространственные структуры 109
6.6. Связь с теорией фазовых переходов Ландау. Вывод уравнений Фоккера–Планка 110

Глава 7. Информация, приращение информации и эффективность самоорганизующихся систем вблизи точек неустойчивости 112

7.1. Введение 112
7.2. Принцип подчинения и его применение к информации . . . 113
7.3. Приращение информации 113
7.4. Пример: неравновесные фазовые переходы 114
7.5. Мягкие одномодовые неустойчивости 116
7.6. Как можно измерить информацию и приращение информации? 116
7.6.1. Эффективность 116
7.6.2. Информация и приращение информации 117
7.7. Случай нескольких параметров порядка 118
7.8. Вычисление информации в случае одного параметра порядка 120
7.8.1. Ниже порога 120
7.8.2. Выше порога 122
7.8.3. Численные результаты 126
7.8.4. Анализ полученных результатов 128
7.9. Точные аналитические результаты относительно информации, приращения информации и эффективности в случае одного параметра порядка 129
7.9.1. Точка неустойчивости 131
7.9.2. Приближение к точке неустойчивости 132
7.9.3. Область устойчивости 133
7.9.4. Введение в систему сигнала 134
7.9.5. Выводы 136
7.10. S-теорема Климонтовича 136
7.10.1. Область 1: значительно ниже порога генерации 138
7.10.2. Область 2: на пороге генерации 139
7.10.3. Область 3: значительно выше порога генерации 140

7.11. Вклад подчиненных мод в информацию вблизи неравновесных фазовых переходов 143

Глава 8. Прямое определение множителей Лагранжа 151

8.1. Информационная энтропия систем ниже и выше их критической точки 151
8.2. Прямое определение множителей Лагранжа ниже критической точки, в критической точке и выше критической точки 153

Глава 9. Непредвзятое моделирование стохастических процессов: как легко и просто получать интегралы по траекториям, уравнения Фоккера–Планка и уравнения Ланжевена–Ито 162

9.1. Одномерный вектор состояния 162
9.2. Обобщение на случай многомерного вектора состояния 165
9.3. Корреляционные функции как ограничения 167
9.4. Уравнение Фоккера–Планка, соответствующее пропагатору на малых временах 170
9.5. Можно ли вывести закон Ньютона из экспериментальных данных? 171

Глава 10. Приложение к некоторым физическим системам 173

10.1. Мультимодовые лазеры с фазовыми соотношениями 173
10.2. Одномодовый лазер с учетом поляризации и инверсии 174
10.3. Гидродинамика: конвективная неустойчивость 176

Глава 11. Переходы между паттернами поведения в биологии. Пример: движения рук 178

11.1. Некоторые экспериментальные факты 178
11.2. Как моделировать переход 180
11.3. Критические флуктуации 186
11.4. Некоторые выводы 192

Глава 12. Распознавание образов 193

12.1. Отбор характерных признаков 193
12.2. Алгоритм распознавания образов 200
12.3. Основные принципы конструирования синергетического компьютера 206
12.4. Обучение с помощью приращения информации 208
12.5. Обучение распознаванию процессов и ассоциативное действие 210

Глава 13. Квантовые системы 216

13.1. Почему возникает необходимость в квантовой теории информации? 216
13.2. Принцип максимума информации 218
13.3. Параметры порядка, подчиненные моды и структуры 225
13.4. Информация, содержащаяся в параметрах порядка и в подчиненных модах 229

Глава 14. Заключительные замечания и некоторые перспективы 231

Литература 234